



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA STROJNÍHO INŽENÝRSTVÍ

FACULTY OF MECHANICAL ENGINEERING

ÚSTAV MECHANIKY TĚLES, MECHATRONIKY A BIOMECHANIKY

INSTITUTE OF SOLID MECHANICS, MECHATRONICS AND BIOMECHANICS

ANALÝZA VLASTNOSTÍ STEREOKAMERY ZED VE VENKOVNÍM PROSTŘEDÍ

ANALYSIS OF ZED STEREOCAMERA IN OUTDOOR ENVIRONMENT

DIPLOMOVÁ PRÁCE

MASTER'S THESIS

AUTOR PRÁCE

AUTHOR

Bc. ONDŘEJ SVOBODA

VEDOUCÍ PRÁCE

SUPERVISOR

doc. Ing. JIŘÍ KREJSA, PhD.

BRNO 2019

Zadání diplomové práce

Ústav: Ústav mechaniky těles, mechatroniky a biomechaniky
Student: **Bc. Ondřej Svoboda**
Studijní program: Aplikované vědy v inženýrství
Studijní obor: Mechatronika
Vedoucí práce: **doc. Ing. Jiří Krejsa, Ph.D.**
Akademický rok: 2018/19

Ředitel ústavu Vám v souladu se zákonem č.111/1998 o vysokých školách a se Studijním a zkušebním řádem VUT v Brně určuje následující téma diplomové práce:

Analýza vlastností stereokamery ZED ve venkovním prostředí

Stručná charakteristika problematiky úkolu:

Stereovize může být v oblasti mobilní robotiky užitečnou a finančně dostupnou alternativou laserových dálkoměrů jako primárního senzoru k získání odhadu polohy robotu a detekci překážek. V posledních letech se zvýšila dostupnost stereokamer nejen pro vnitřní prostředí, ale také pro prostředí vnější, s řádově vyšším dosahem, tedy maximální detekovatelnou vzdáleností překážky. Cílem práce je analyzovat použitelnost konkrétní stereokamery ZED pro její použití v mobilní robotice.

Cíle diplomové práce:

Zprovozníte stereokameru ZED v prostředí ROS.

Navrhnete metodiku sběru dat pro pozdější offline analýzu vlivu veličin podstatných z hlediska kvality detekované hloubkové mapy.

Provedte sérii měření ve venkovním prostředí, například parku.

Analyzujte vliv veličin podstatných pro kvalitu výsledné hloubkové mapy, například osvětlení scény, rychlost pohybu kamery.

Vyberte, implementujte a verifikujte algoritmus detekce orientačních bodů v získaných hloubkových mapách.

Seznam doporučené literatury:

SDK kamery ZED, dostupné online na adrese „[stereolabs.com](https://www.stereolabs.com)“

Pedro NÚNEZ, Ricardo VÁZQUEZ-MARTÍN, Antonio BANDERA, Visual Odometry Based on Structural Matching of Local Invariant Features Using Stereo Camera Sensor, Sensors 2011, 11(7), 7262-7284; doi:10.3390/s110707262

Termín odevzdání diplomové práce je stanoven časovým plánem akademického roku 2018/19

V Brně, dne

L. S.

prof. Ing. Jindřich Petruška, CSc.
ředitel ústavu

doc. Ing. Jaroslav Katolický, Ph.D.
děkan fakulty

Abstrakt

Tato diplomová práce je zaměřena na testování ZED kamery a SLAM mapování ve venkovním prostředí. Je zde porovnávána funkčnost samotné vizuální odometrie ZEDfu, vyvíjena samotným výrobcem kamery, s běžně používanými metodami pro sledování trajektorie a to pomocí GPS nebo kolové odometrie. Dále se zde testuje SLAM mapování v RTAB-Map v závislosti na proměnných podmínkách prostředí a to za použití dvou metod BRISK a SIFT. Provedená analýza by měla sloužit pro pozdější aplikace ZED kamery v mobilní robotice.

Summary

The Master thesis is focused on analyzing stereo camera ZED in the outdoor environment. There is compared ZEDfu visual odometry with commonly used methods like GPS or wheel odometry. Moreover, the thesis includes analyses of SLAM in the changeable outdoor environment, too. The simultaneous mapping and localization in RTAB-Map were processed separately with SIFT and BRISK descriptors. The aim of this master thesis is to analyze the behaviour ZED camera in the outdoor environment for future implementation in mobile robotics.

Klíčová slova

BRISK, epipolární geometrie, FLANN, GPS, kolová odometrie, mobilní roboti, příznakové metody, RANSAC, ROS, RTAB-Map, SIFT, simultánní lokalizace a mapování (SLAM), Stereolabs, stereokamera ZED, V-SLAM , vizuální odometrie, venkovní prostředí

Keywords

BRISK, computer vision, epipolar geometry, FLANN, GPS, motion estimation, outdoor environment, RANSAC, ROS, RTAB-Map, SIFT, simultaneous localization and mapping (SLAM), stereo camera ZED, Stereolabs, V-SLAM, visual odometry, wheel odometry

SVOBODA, O. *Analýza vlastností stereokamery ZED ve venkovním prostředí*. Brno: Vysoké učení technické v Brně, Fakulta strojního inženýrství, 2019. 58 s. Vedoucí diplomové práce doc. Ing. Jiří Krejsa, PhD.

Prohlašuji, že jsem diplomovou práci na téma „Analýza vlastností stereokamery ZED ve venkovním prostředí“ vypracoval samostatně s použitím odborné literatury a pramenů uvedených v seznamu použitých zdrojů.

V Brně dne 22. května 2019

Bc. Ondřej Svoboda

Děkuji svému vedoucímu doc. Ing. Jiřímu Krejsovi, PhD. za vedení mé diplomové práce a veškeré poskytnuté prostředky při její tvorbě. Také děkuji své rodině za poskytnutou podporu během mého studia.

Zvláštní poděkování za výborné lektvary moudrosti, bez kterých by bylo těžké dokončit studium, patří také mocnému mágovi Tomáši Máchovi.

Bc. Ondřej Svoboda

Obsah

1	Základy vizuální odometrie	4
1.1	Úvod do vizuální odometrie	4
1.1.1	Historie vizuální odometrie	4
1.2	Dělení metod ve vizuální odometrii	5
2	Matematický model stereokamery	7
2.1	Projekční matice	7
2.2	Epipolární geometrie	8
2.3	Určení bodu z pixelových souřadnic v obraze	10
2.4	Kalibrace kamery	12
2.4.1	Prvky v kalibrační matici	12
3	Příznakové metody	14
3.1	Metoda 2D do 2D	14
3.1.1	Longuet-Higginsova osmibodová metoda	15
3.1.2	Určení matice rotace a translace z esenciální matice	16
3.2	Metoda 3D do 3D	17
3.3	Metoda 3D do 2D	18
3.3.1	Metoda P6P	19
4	Detektory a deskriptory příznaků	20
4.1	SIFT	20
4.2	BRISK	22
5	Přiřazování a filtrování příznaků	25
5.1	Přiřazování příznaků	25
5.1.1	Knihovna FLANN	25
5.2	Filtrování příznaků	26
5.2.1	RANSAC	26
6	RTAB-Map	27
6.1	Synchronizace dat	27
6.2	STM	28
6.3	Detekce uzavřených smyček a přibližná lokalizace	30
6.4	Optimalizace grafu	30
6.5	Vytváření globální mapy	31
7	Návrh analýzy a vyhodnocení dat	32
7.1	Návrh analýzy dat	33
7.2	Mapování	34
7.3	Lokalizace	43
7.3.1	Lokalizace za použití metody BRISK	44
7.3.2	Lokalizace za použití metody SIFT	47
7.3.3	Srovnání lokalizace	49

Úvod

Vizuální odometrie je jeden ze způsobů zpracování detekce pohybu robota z jedné nebo z více kamer [1]. Název „vizuální odometrie“ byl zvolený na základě podobnosti kolové odometrie. Rozdíl spočívá v tom, že pohyb není určen na základě integrujícího se počtu otáček kol, nýbrž na detekci pohybu obrazu kamery. Obecně je také součástí problematiky zvané SFM (Structure From Motion), která se zabývá polohou kamery a vytvářením 3D prostoru ze snímaných obrazových dat. V současné době se vizuální odometrie využívá v mnoha oblastech například v navigaci robotů, v autonomních systémech či ve vesmírném výzkumu, kde je nedílnou součástí výzkumných robotických vozítek.

Tato diplomová je členěna do několika částí. V úvodní části je pojednáno o vizuální odometrii, o její historii a především jejich nezbytných základech. V dalších částech jsou vysvětleny metody, kterými je zpracován pohyb na základě zpracování obrazových dat, samotné korekce, filtrace a v případě lokalizace i typy vyhledávání příznaků. Pro tzv. simultánní lokalizaci a mapování (SLAM) je použit program RTAB-Map vyvíjen v prostředí ROS. Popis funkce tohoto programu včetně technických požadavků a parametrů doplňuje teoretickou část této diplomové práce. V praktické části je popsána samotná analýza, mapování a lokalizace s jejími výsledky. Jelikož je česká terminologie nejednotná, vyskytují se v této diplomové práci pojmy v anglickém jazyce.

Počítačové vidění patří významná dynamicky vyvíjející se odvětví výpočetní techniky a nachází uplatnění v mnoha oblastech. Do tohoto oboru se řadí i prostorové vidění tzv. stereovize. Tento způsob techniky se používá zejména v oblasti mobilní robotiky, kde je využívána pro výpočet odometrie a detekci překážek. Jsou finančně dostupné a levné alternativy pro určování polohy a detekci překážek ve vnitřním a venkovním prostředí. Výhodou vizuální odometrie je, že lze pomocí aplikování SLAM (simultánní lokalizace a mapování) eliminovat chyby vzniklé při zpracování obrazových dat.

Stereokamera ZED je vyvíjena společností Stereolabs a je využívána nejenom v mobilní robotice, ale i při vytváření virtuální reality. K produktu jsou výrobcem dodány přesné kalibrační parametry, software a knihovny určené pro různé vývojářské aplikace včetně vizuální odometrie ZEDfu vyvíjenou výrobcem kamery. Je použitelná i v různých prostředích jako MATLAB, ROS, ROS2, OpenCV nebo TensorFlow. Cílem této práce je analyzovat její funkčnost včetně SLAM v proměnných podmínkách ve venkovním prostředí pro její použití v mobilní robotice.

1. Základy vizuální odometrie

1.1. Úvod do vizuální odometrie

Vizuální odometrie je proces určování pohybu určitého objektu (například vozidla, člověka nebo robota) na základě obrazových vstupů z jedné či více kamer. Při pohybu daného objektu dochází k určitým změnám, při kterých se obrazové vstupy mění a jejichmi změnami se určuje přibližný pohyb a přibližná poloha sledovaného objektu.

1.1.1. Historie vizuální odometrie

Počátky vizuální odometrie sahá do 80. let minulého století, kde první zmínky o vizuální odometrii lze dohledat v dizertační práci Hanse Moravce [2]. Její vývoj byl financována americkým národním úřadem pro letectví a kosmonautiku NASA [3],[4],[5],[6]. Vizuální odometrie byla vyvíjena pro planetární moduly na Marsu, aby byly schopny měřit pohyb se šesti stupni volnosti nezávisle na terénu, ve kterém by mohlo dojít k prokluzování kol a tím i k velké nepřesnosti měření. Běžná kolová odometrie, u které je určován pohyb z enkodérů, by tedy byla zcela nepoužitelná. Termín „vizuální odometrie“ byl použit teprve až v roce 2004 ve stejnojmenném článku [7].

První prototypy, u kterých byla implementována vizuální odometrie, nepracovaly plynule, ale krokově. Navíc byly prakticky nepoužitelné kvůli výpočtové náročnosti, které tehdejší procesory výkonově nezvládaly. V první použité metodě od Hanse Moravce se po každém kroku analyzovalo devět fotografií v určitých časových okamžicích. Kamera popojížděla na nosníku a ve stejných vzdálenostech pořídila devět fotografií. Z těch byly vybrány jisté příznaky a byly k sobě přiřazeny v dalších osmi snímcích použitím vzájemné korelace. Při další změně polohy bylo opět pořízeno devět snímků a potenciální body byly nalezeny vzájemnou korelací, která byla založená na „coarse-to fine“ strategii. Nežádoucí příznaky byly odstraněny hloubkovou analýzou v osmi stereo párech.

V konečné fázi byla spočítána poloha a to transformací prostorového objektu ze dvou po sobě následujících poloh objektu.

Při tomto výpočetním procesu vznikalo nespočet rovnic, které byly následně spočítány metodou WLS (Weighted Least Square). Váha jednotlivých rovnic v této metodě závisí na vzdálenosti bodu.

Jak již bylo zmíněno, v této metodě nebyla zahrnuta stereokamera, nýbrž jen jedna kamera popojíždějící po příčce. I přesto se tato metodika se řadí ke stereo vizuální odometrii. Alternativou stereo vizuální odometrie je i monokulární vizuální odometrie, tedy vizuální odometrie s jednou kamerou. V tomto případě vzniká nevýhoda v tom, že je závislá na škálovacím faktoru a při zpracování dat jsou nutné i data z ostatních senzorů (např. IMU) narozdíl od stereo vizuální odometrie k dosažení absolutního škálování. Pokud se stane, že vzdálenost mezi kamerami je nesrovnatelně menší než vzdálenosti v obraze, metodu stereo vizuální odometrie již pak nelze nadále používat a je třeba přejít na monokulární vizuální odometrii.

Vizuální odometrie prošla od roku 1980 bouřlivým vývojem a prvního reálného imple-

mentování se dočkala až po roce 2000, kdy byla úspěšně implementována do mobilních robotických vozítek (roverů) na Marsu.

Vizuální simultánní lokalizace a mapování (V-SLAM)

V-SLAM (Visual Simultaneous Localization and Mapping) je simultánní lokalizace a mapování, kde kromě k výpočtu odometrie dochází i k vytváření mapy prostředí. Cílem této metody je dosáhnout globálního konzistentního mapování trajektorie, přičemž v průběhu snímání prostředí vytváří vlastní trajektorii. V průběhu tohoto procesu dochází i k akumulaci chyb, kterými není možné určit přesnou trajektorii. Tyto chyby lze značně eliminovat. Pokud systém rozezná, že se objekt vrátil na některé z původních míst, dojde k uzavření smyčky a poté je trajektorie přepočítána tak, aby tuto smyčku obsahovala. V tom se zásadně liší od samotné vizuální odometrie, která se soustředí pouze na lokální mapování, zatímco V-SLAM se zabývá globálním mapováním. Je také preciznější a v praxi je více používanější než samotná vizuální odometrie. Nevýhodou V-SLAM je ovšem to, že uzavřené smyčky mohou negativně ovlivnit mapování a navíc implementace této metody je z hlediska výpočetního výkonu a dalších technických nároků daleko náročnější [1].

1.2. Dělení metod ve vizuální odometrii

Vizuální odometrie je možné členit do několika následujících skupin:

1. Podle počtu použitých kamer

(a) Monokulární metody

Monokulární odometrie je vizuální odometrie při použití jedné kamery. Prostorová projekce je přepočítávána z 2D dat. Z počátku je vzdálenost mezi prvními dvěma pozicemi nastavena na hodnotu jedné. V následujících krocích se pak mění relativní škálovací odchylka na základě jiných naměřených dat.

(b) Stereo metody

Metody využívající stereo metody využívají dvě kamery (stereokamera). Pomocí nich dokáží přesně určit polohou snímaných objektů v prostoru. Také odhadují i posun mezi snímky.

(c) Vícekamerové metody

Vícekamerové nebo-li multikamerové, využívají tři a více kamer. Takovým případem může být například V-SLAM typu MCPTAM nebo RTAB-Map.

2. Podle typu rozeznávání

Změnu mezi jednotlivými snímky lze určovat pomocí příznaků (features). Ty mohou být ve formě bodů či různých linií. Těmto bodům bývají přiřazeny jednoznačné identifikátory (deskriptory). Metody, kterými jsou tyto význačné body rozeznávány, lze dělit do dvou skupin a to na příznakové a globální.

(a) Globální metody

Globální metody jsou starší a posunutí obrazu se určuje na základě vzájemné korelace mezi dvěma snímky. Tato metoda se již takřka nevyužívá.

1.2. DĚLENÍ METOD VE VIZUÁLNÍ ODOMETRII

(b) Příznakové metody

Příznakové metody jsou založeny na hledání společných příznaků mezi snímky. Výstupem této metody je množina korespondujících bodů a následně pomocí nich se určuje posun. Tyto metody jsou robustnější a na výpočet méně náročné. Používá se ve všech dostupných metodách vizuální odometrie.

Příznakové metody se dále dělí podle typu transformace T_k dvou po sobě jdoucích snímcích.

i. *Metoda 2D do 2D*

Tato metoda se vyznačuje tím, že oba korespondující příznaky jsou specifikované v rovinných souřadnicích. Slouží především pro monokulární kamery.

ii. *Metoda 3D do 3D*

V této metodě korespondující příznaky f_{k-1} v předchozím obraze a příznaky f_k v současném obraze jsou charakterizovány v prostorových souřadnicích. Proto, aby mohly být tyto souřadnice definovány, je zcela nezbytná prostorová bodová triangulace prostoru v každém časovém okamžiku. Tato metoda je využívána u starších typů vizuální odometrie.

iii. *Metoda 3D do 2D*

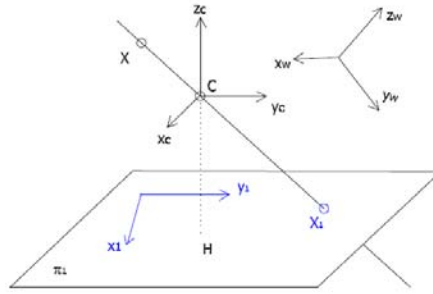
Tato transformace patří mezi dosud nejpřesnější a nejpoužívanější. V tom případě jsou příznaky f_{k-1} v předchozím obraze v prostorových souřadnicích a příznaky v současném obraze jsou f_k jsou charakterizovány v rovinných souřadnicích. V případě monokulární kamery je 3D struktura triangulována ze dvou předchozích snímků a poté jsou příznaky přiřazeny k příznakům současného obrazu, které jsou ve 2D souřadnicích.

2. Matematický model stereokamery

Jelikož není možné pouze za použití jedné kamery určit polohu bodu v prostoru, je proto nutné více kamer. Nejběžnějším případem je aplikace stereokamery s dvěma zabudovanými kamerami s různými středy promítání. Matematický model představuje epipolární geometrie sestavená ze dvou perspektivních modelů kamery. Jelikož v tomto procesu hraje významnou roli i vztah mezi pixelovými souřadnicemi snímku a souřadnicemi pozorovaného objektu v pevně daných afinních souřadnic systému (repéru), je tato transformace určena projekční maticí kamery.

2.1. Projekční matice

U geometrického modelu kamery se předpokládá, se prostorová scéna středově promítá ze středu C na průmětnu π . V systému se nachází tři souřadné systémy a to pevně daný repér v obecné poloze vůči průmětně $W < O_w, x_w, y_w, z_w >$, dále repér objektivu kamery, kde se souřadná osa z shoduje s optickou osou $C < C, x_c, y_c, z_c >$ a posledně jsou to ortogonální obrazové (pixelové) souřadnice obrázku $I < O, x_1, y_1 >$ [8]. Dále se v na obrázku nachází pozorovaný bod X .



Obrázek 2.1: Projekční matice kamery [8]

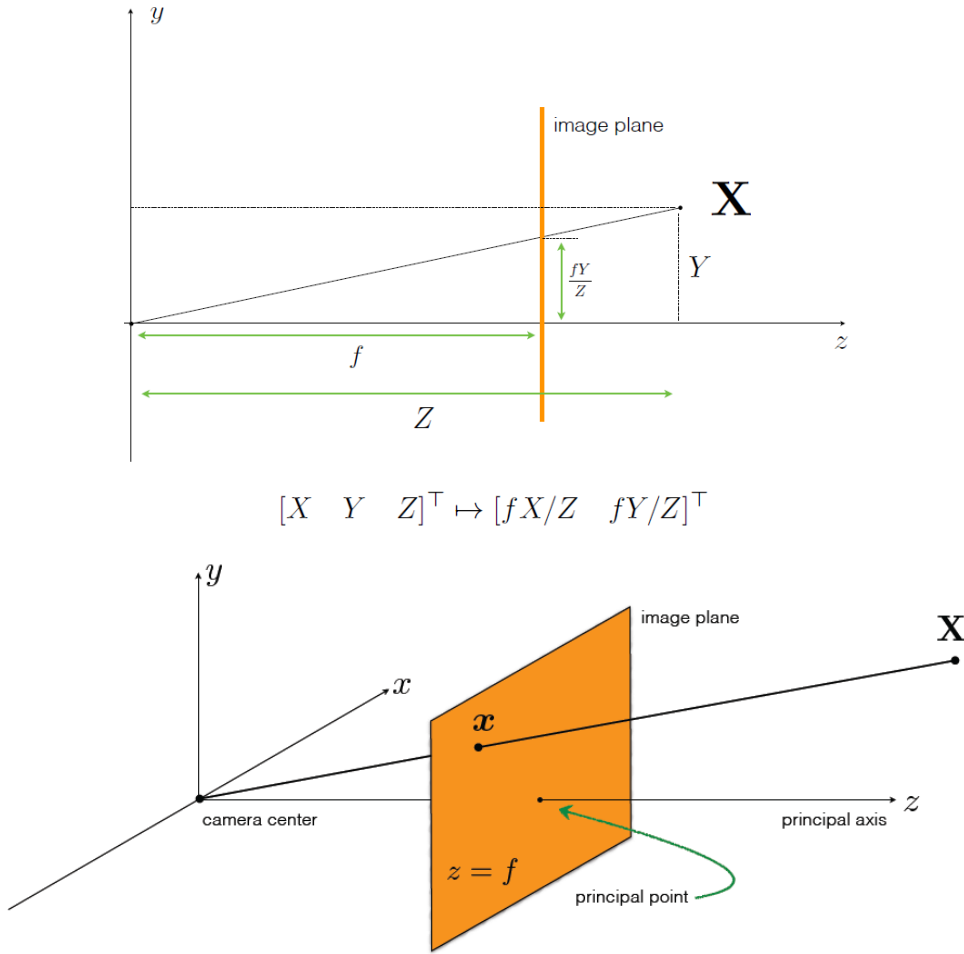
U středového promítání, kde se obecný bod zobrazuje do promítací roviny π_1 $X \rightarrow X_1; (w, x, y, z)^T \rightarrow (w_1, x_1, y_1, z_1)^T$ platí následující matice [8]:

$$P_{persp} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & -f & 1 \end{pmatrix} = (0\tilde{P}_{persp}) \quad (2.1)$$

Parametr f je dle obrázku 2.2 vzdálenost středu od průmětny. S obrazovými informacemi o obrazových souřadnicích se také ukládá pět tzv. vnitřních parametrů kamery, mezi které kromě parametru f , se taktéž řadí parametry určující posunutí počátku, změnu měřítka ve směru os a afinní transformaci do neortogonální báze obrázku.

Transformace mezi pevně daným repérem W a repérem kamery C je definována tzv. vnějšími parametry a to posunutím \vec{t} a ortogonální maticí rotace $R \in O(3, R)$. Konečná projekční matice P je složena z vnitřní kalibrace kamery, z matice perspektivní projekce a z vnější kalibrace kamery [8]:

2.2. EPIPOLÁRNÍ GEOMETRIE



Obrázek 2.2: Znázornění perspektivní projekce [9]

$$\begin{aligned}
 P &= \begin{pmatrix} 1 & 0 & 0 \\ x_0 & a & b \\ y_0 & 0 & c \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & -f & 0 & 0 \\ 0 & 0 & -f & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \vec{t} & R \end{pmatrix} = \\
 &\begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & -af & -bf & x_0 \\ 0 & 0 & -cf & y_0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \vec{t} & R \end{pmatrix} = (0K) \begin{pmatrix} 1 & 0 \\ \vec{t} & R \end{pmatrix} = K(\vec{t} R)
 \end{aligned} \tag{2.2}$$

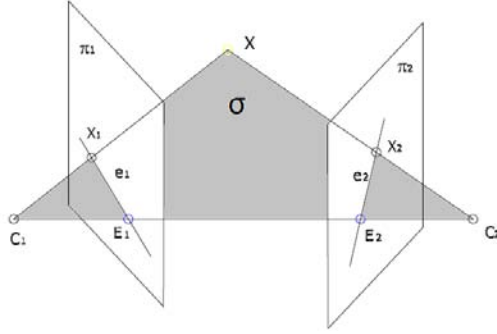
Matice K je typu 3×3 je nazývána jako kalibrace kamery. Pro transformaci souřadnic mezi bodem v prostoru a jeho obrazu v projektivní rovině platí následující rovnice [8]:

$$X_1 = P \cdot X; \quad P = K(\vec{t} R) \tag{2.3}$$

2.2. Epipolární geometrie

Matematický model stereokamery se skládá ze dvou projekčních rovin dvou kamer o dvou různých kalibracích K_1 a K_2 . Rovina σ je charakterizována třemi body a to body C_1 , C_2 a obecným bodem X . Promítací paprsky se nacházejí v rovině σ a podle obrázku

2. MATEMATICKÝ MODEL STEREOKAMERY



Obrázek 2.3: Nákres epipolární geometrie [8]

2.3 se všechny promítací paprsky nacházejí na sdružených přímkách (epipolárách), což jsou průsečnice dvou rovin π_1 a π_1 . Průsečnicí (epipolárou) e_1 mezi rovinou σ a projekční rovinou π_1 od jedné kamery a druhou průsečnicí (epipolárou) e_2 mezi rovinou σ projekční rovinou π_2 od druhé kamery. U matematických rovnic 2.4 a 2.5, charakterizujících vztahy mezi průměty X_1 , X_2 a X , se předpokládá, že pevný souřadný systém je totožný se souřadným systémem jedné z kamer. Vztah mezi průměty bodů a obecným bodem je charakterizován v rovnicích 2.4 a 2.5 [8]:

$$X_1 = P_1 X = K_1(0, I)X \quad (2.4)$$

$$X_2 = P_2 X = K_2(\vec{t} R)X \quad (2.5)$$

Vektor \vec{t} a matice rotace R popisují transformace mezi objektivy kamer. Vztahy mezi průměty a obecným bodem lze definovat i pomocí normalizovaných souřadnic \hat{X}_1 a \hat{X}_2 [8].

$$X_1 = K_1 \hat{X}_1 \quad (2.6)$$

$$X_2 = K_2 \hat{X}_2 \quad (2.7)$$

Epipolární geometrie je popsána dvěma maticemi a to fundamentální a esenciální o rozměrech 3×3 a hodnoti 2. Esenciální matice [10] je charakterizována rotační maticí R a antisymetrickou maticí vektoru posunutí t_M o rozměrech 3×3 .

$$E = t_M R \quad (2.8)$$

Vektorový součin mezi normalizovanými obrazovými souřadnicemi a esenciální maticí je roven nule [8].

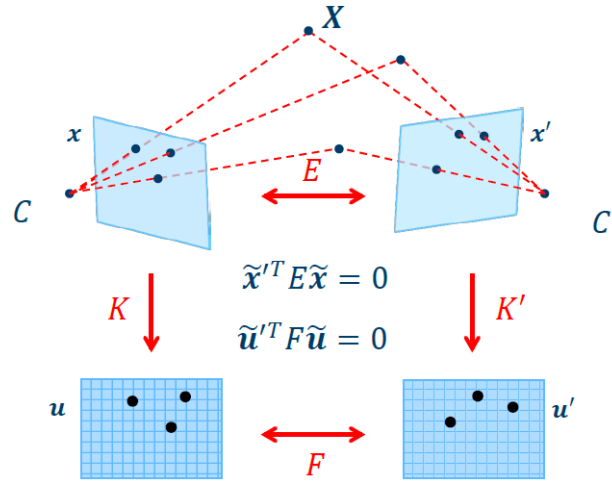
$$\hat{X}_2^T \cdot t_M R \cdot \hat{X}_2 = 0 \quad (2.9)$$

Dosazením rovnic 2.6 a 2.7 do rovnice 2.9 se získá fundamentální matice, která má tvar $F = K_1^{-T} E K_2^{-1}$.

$$X_2^T (K_1^{-T} E K_2^{-1}) X_1 = 0 \quad (2.10)$$

Fundamentální matice F je jednoznačně charakterizována dvěma snímky a tato matice určuje dvojici projekčních matic kamer P_1 a P_2 . Pro její výpočet je nezbytné určit sedm parametrů, což je zásadním úkolem tzv. stereo analýzy. Rozdíl mezi fundamentální maticí F a esenciální maticí E je znázornění na obrázku 2.4.

2.3. URČENÍ BODU Z PÍXELOVÝCH SOUŘADNIC V OBRAZE

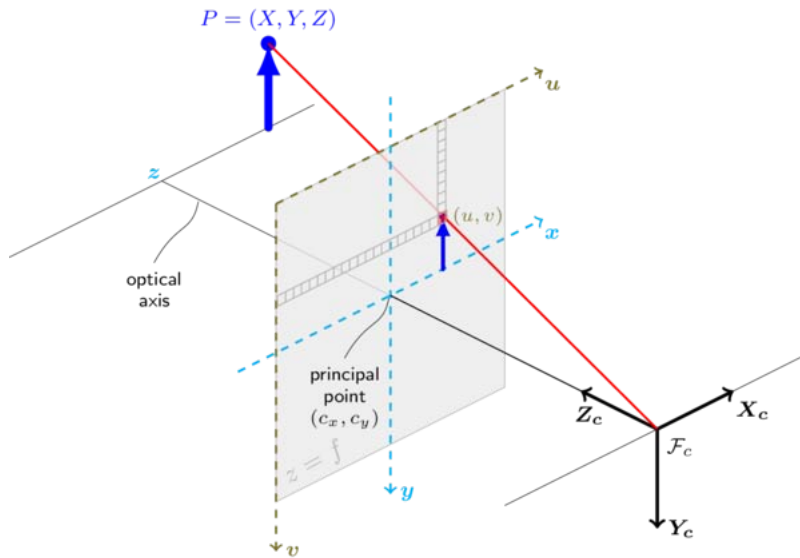


Obrázek 2.4: Esenciální a fundametrální matice[11]

2.3. Určení bodu z pixelových souřadnic v obraze

Pro určení bodů v 3D prostoru pomocí stereokamery je tedy nezbytné znát kalibrační matice K_1 a K_2 , dále matice složená z translace \vec{t} a matice rotace R . Vše je tedy zahrnuto do matic P_1 a P_2 , které slouží k transformaci do prostorových souřadnic. Podle rovnic 2.4 a 2.5 vypadají matice následovně:

$$P_1 = K_1 \cdot \begin{bmatrix} 1 & 0 & 0 & | & 0 \\ 0 & 1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \end{bmatrix}, \quad P_2 = K_2 \cdot [R | \vec{t}] \quad (2.11)$$



Obrázek 2.5: Perspektivní model kamery [12]

Podle obrázku 2.5 lze rovnice 2.11 přepsat do následujícího tvaru, kde na levé straně vektor (u, v) představuje vektor charakterizující pozici bodu v promítací rovině násobenou volným parametrem λ .

$$\lambda \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = P \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} \vec{p}_1^T \\ \vec{p}_2^T \\ \vec{p}_3^T \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad (2.12)$$

Jelikož při převodu z pixelových na globální souřadnice vznikají algebraické chyby, je třeba tyto chyby minimalizovat. Z rovnice 2.12 je zřejmé, že mezi pixelovými souřadnicemi a globálními souřadnicemi je lineární závislost. Proto musí platit, že vektorový součin musí být roven nule. V následujícím odvození rovnice 2.13 citovanou ze zdroje [11] jsou \vec{p}_1^T , \vec{p}_2^T a \vec{p}_3^T řádkové vektory. \tilde{X}_i a \tilde{u}_i představují difference.

$$\begin{aligned} \tilde{u}_i &= P\tilde{X}_i \\ \Downarrow \\ \tilde{u}_i \times P\tilde{X}_i &= 0 \\ \Downarrow \\ \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \times \begin{bmatrix} \vec{p}_1^T \\ \vec{p}_2^T \\ \vec{p}_3^T \end{bmatrix} \tilde{X}_i &= 0 \\ \Downarrow \\ \begin{bmatrix} v_i \vec{p}_3^T - \vec{p}_2^T \\ \vec{p}_1^T - u_i \vec{p}_3^T \\ u_i \vec{p}_2^T - v_i \vec{p}_1^T \end{bmatrix} \tilde{X}_i &= 0 \\ \Updownarrow \\ \begin{bmatrix} v_i \vec{p}_3^T - \vec{p}_2^T \\ u_i \vec{p}_3^T - \vec{p}_1^T \end{bmatrix} \tilde{X}_i &= 0 \end{aligned} \quad (2.13)$$

Lineární triangulaci [13] lze tedy zapsat v následujícím tvaru, kde vektory u a v jsou obrazové souřadnice první kamery, u' a v' jsou obrazové souřadnice druhé kamery. Vstupem jsou čtyřdimenzionální vektory \vec{X}_{4D} , \vec{Y}_{4D} , \vec{Z}_{4D} , \vec{W}_{4D} , představující homogenní souřadnice bodů \vec{X} , \vec{Y} a \vec{Z} [14].

$$\begin{bmatrix} u_i \vec{p}_3^T - \vec{p}_1^T \\ v_i \vec{p}_3^T - \vec{p}_2^T \\ u'_i \vec{p}_3^T - \vec{p}_1^T \\ v'_i \vec{p}_3^T - \vec{p}_2^T \end{bmatrix} \begin{bmatrix} \vec{X}_{4D} \\ \vec{Y}_{4D} \\ \vec{Z}_{4D} \\ \vec{W}_{4D} \end{bmatrix} = A \cdot \begin{bmatrix} \vec{X}_{4D} \\ \vec{Y}_{4D} \\ \vec{Z}_{4D} \\ \vec{W}_{4D} \end{bmatrix} = \vec{0} \quad (2.14)$$

Čtyřdimenzionální vektor je homogenní souřadnice bodu \vec{X} , \vec{Y} a \vec{Z} vynásobenou skalárem. Souřadnici lze získat vydělením posledním prvkem \vec{W}_{4D} [14].

$$X = \frac{X_{4D}}{W_{4D}}, \quad Y = \frac{Y_{4D}}{W_{4D}}, \quad Z = \frac{Z_{4D}}{W_{4D}} \quad (2.15)$$

2.4. KALIBRACE KAMERY

Určené souřadnice z kamery jsou zatíženy chybou, a tudíž je velmi problematické počítat pozici bodu pouhým dosazením. Proto je nutné při výpočtu používat již předurčené soustavy rovnic.

2.4. Kalibrace kamery

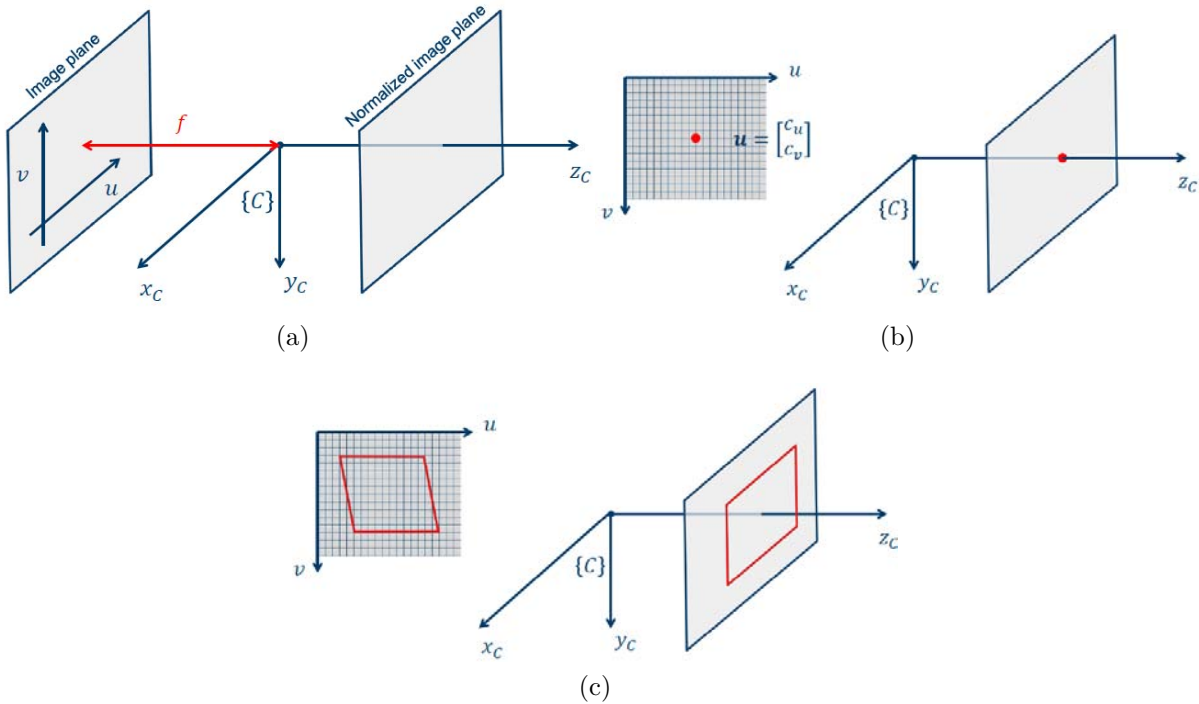
Při správném chodu stereo analýzy je nezbytné znát kalibrační parametry určených ke korekci obrazového systému. Kromě těchto dalších parametrů v kalibrační matici se určují parametry, které korigují vady čočky.

2.4.1. Prvky v kalibrační matici

V předchozí kapitole je odvozena kalibrační matice 2.2, která obsahuje prvky a , b , c , f , x_0 a y_0 . Po ekvivalentní řádkové úpravě této kalibrační matice se získá stejný tvar matice stejný tvar jako v článku [15].

$$K = \begin{bmatrix} 0 & 0 & 1 \\ -af & -bf & x_0 \\ 0 & -cf & y_0 \end{bmatrix} \sim \begin{bmatrix} -af & -bf & x_0 \\ 0 & -cf & y_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_u & s & c_u \\ 0 & f_v & c_v \\ 0 & 0 & 1 \end{bmatrix} \quad (2.16)$$

Parametry f_u a f_v představují vzdálenosti projekční roviny kamery. Prvky c_u a c_v představují body průsečíku osy Z a projekční roviny. Poslední prvek s představuje je korekční parametr, který upravuje zkosení zobrazení čočky kamery.



Obrázek 2.6: (a) Znázornění parametru f [15], (b) Znázornění parametru c [15], (c) Znázornění vlivu parametru s [15]

Teoreticky je sice možné tyto parametry dosadit přímo do matic na základě určených parametrů kamer, ale v praxi není možné tyto parametry určit s velikou přesností, jakou je

2. MATEMATICKÝ MODEL STEREOKAMERY

nutné dodržet. Proto jsou obvykle určeny metodami pro kalibraci kamer, které určují tyto parametry z obrazových informací. Obvykle tyto metody probíhají za použití speciálních značek, pomocí kterých se dopočítávají jednotlivé parametry do kalibrační matice.

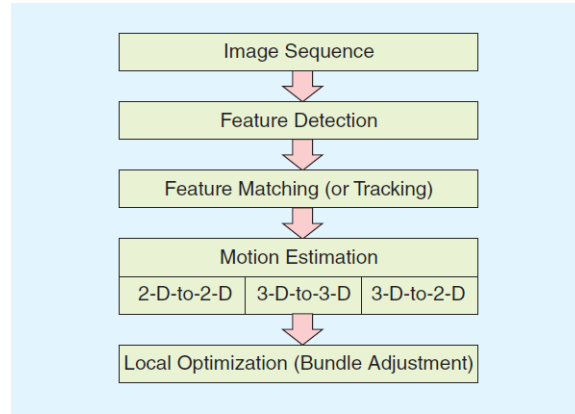
U ZED stereokamery je možnost tyto parametry získat z internetových stránek výrobce, popřípadě je také možné použít software od výrobce. Software pro kalibraci ZED kamery využívá šachovnice, na kterých se zobrazují oblasti, kam uživatel musí stereo kameru zamířit. Kromě parametrů do kalibrační matice se pomocí této metody určují i distorzní koeficienty, jejichž počet se pohybuje kolem osmi. Dva až šest distorzních koeficientů pro radiální zkreslení a další dva distorzní koeficienty pro tangenciální zkreslení. Při této kalibraci je součástí i určení i fundamentální matice, esenciální matice, rotační matice a vektoru posunutí.

3. Příznakové metody

Cílem vizuální odometrie je určit trajektorii pohybu na základě změny vstupních obrazů. Obecně je tato změna charakterizována změnou rotační a translační matice. Do řídicího systému vstupují snímky, ze kterých různými metodami vybírají příznaky. Z korespondujících příznaků z minulých a současných snímků se určí grafické transformace, tedy rotace a posunutí. U stereokamery je posun určován zpravidla z předešlého a současného snímku. Ve výjimečném případě se určuje i ze dvou posledních snímků. Maticově lze změna pohybu zapsat následující maticí [14]:

$$T_{i,i-1} = \begin{bmatrix} R_{i,i-1} & \vec{t}_{i,i-1} \\ 0 & 1 \end{bmatrix} \quad (3.1)$$

Index i značí pořadí snímku, T určuje relativní posun, R je rotační matice a \vec{t} je translační vektor. Blokový diagram celého procesu lze obecně pospat následujícím blokovým diagramem [1]:



Obrázek 3.1: Blokový diagram vizuální odometrie [1]

Do blokového diagramu vstupují obrazy. V následující části se detekují příznaky, které jsou poté přiřazovány nebo sledovány. V dalším kroku dochází ke spárování korespondující si dvojice bodů. Pro vyšší přesnost se tento příznak prohledává pomocí nástrojů pro lokální prohledávání. Posunutí a rotace jsou určeny na základě změny pozice korespondujících bodů. Výpočty posunutí a rotace z navzájem si korespondujících bodů je vypočítáno metodami z *2D do 2D*, *3D do 3D* nebo v současné době nepoužívanější metodou *3D do 2D*.

3.1. Metoda 2D do 2D

Posun mezi dvěma snímky I_k a I_{k-1} kalibrovanou kamerou je definován výpočtem esenciální matice. Esenciální matice E je závislá na translačním vektoru \hat{t}_k a matice rotace R_k . Translační vektor nemá jednoznačně určené měřítko, nýbrž je jen relativní. Esenciální matici lze charakterizovat následující rovnicí [1]:

$$E_k \simeq \hat{t}_k R_k \quad (3.2)$$

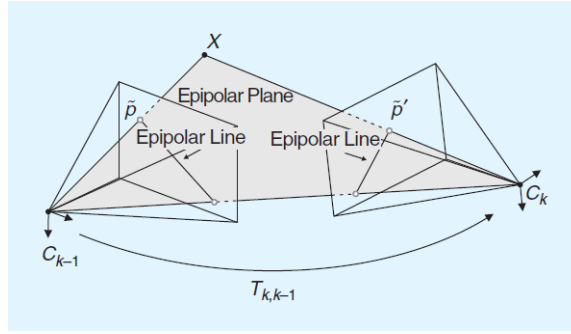
$t_k = [t_x, t_y, t_z]^\top$ je translační vektor, jehož normalizovaný vektor je charakterizován následující maticí.

$$\hat{t}_k = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \quad (3.3)$$

Esenciální matice je určena příznaky určených pouze v rovinných souřadnicích. Rotace a translace je přímo určena z esenciální matice E . Body \tilde{p} a \tilde{p}' představují normalizované body průsečíků epipolárních linií v projekční rovině, které jsou formulovány $\tilde{p} = [\tilde{u}, \tilde{v}, 1]^\top$ podle následujícího obrázku 3.2.

Vztah mezi obrazovými vektory a esenciální maticí je charakterizován následujícím vztahem [1]

$$\tilde{p}'^T E \tilde{p} = 0 \quad (3.4)$$



Obrázek 3.2: Epipolární geometrie pro metodu z 2D do 2D [1]

Esenciální matici lze určit pomocí dvou algoritmů. Prvním algoritmem je Niesterova pětibodová metoda [16], což je minimální počet korespondencí pro určení esenciální matice. Výhodou této metody je, že funguje i pro body ležící v jedné rovině, ale naopak vyžaduje přesnou kalibraci kamery. Jednoduchou a jednoznačnou metodou je ta, která využívá osm a více korespondencí. Jednou z nich je Longuet-Higginsova osmibodová metoda [17] používaná v RTAB-Map.

3.1.1. Longuet-Higginsova osmibodová metoda

Longuet-Higginsovu osmibodovou metodu lze odvodit tímto způsobem. Rovnici 3.4 lze upravit do následujícího tvaru

$$\tilde{p}'^T E \tilde{p} = 0 \Rightarrow [\tilde{u}' \quad \tilde{v}' \quad 1] \cdot \begin{bmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{bmatrix} \cdot \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ 1 \end{bmatrix} = 0 \quad (3.5)$$

Vektor E je v následujícím složení:

$$E = [e_1 \quad e_2 \quad e_3 \quad e_4 \quad e_5 \quad e_6 \quad e_7 \quad e_8 \quad e_9]^\top \quad (3.6)$$

Roznásobením a úpravou rovnice 3.5 lze získat následující tvar důležitý pro výpočet esenciální rovnice:

$$[\tilde{u}\tilde{u}' \quad \tilde{u}'\tilde{v} \quad \tilde{u}' \quad \tilde{u}\tilde{v}' \quad \tilde{v}\tilde{v}' \quad \tilde{v}' \quad \tilde{u} \quad \tilde{v} \quad 1] E_C = 0 \quad (3.7)$$

3.1. METODA 2D DO 2D

Dosazením osmi korespondujících bodů do rovnice 3.7 se získá matice A o velikosti 8×9 . Jejím maticovým násobením s vektorem E vznikne homogenní soustava lineárních rovnic, která má tvar [14]:

$$A \cdot E_C = \vec{0} \quad (3.8)$$

Dosazením osmi korespondujících bodů do rovnice 3.7 vznikne soustava lineárních homogenních lineárních rovnic. Při řešení této soustavy se využívá metoda singulárního rozkladu matice (SVD). Matice A se podle SVD rozkládá do třech matic U , S a V [1].

$$A = U \cdot S \cdot V^T \quad (3.9)$$

Prvky parametru E se při této metodě řešení nachází v posledním sloupci matice V , pro které platí následující rovnice:

$$||E_C|| = 1 \quad (3.10)$$

Tyto prvky parametru E_C neodpovídají prvkům esenciální matice E . U esenciální matice platí, že první dvě vlastní čísla jsou stejná a třetí je nulové. Proto při singulárním rozkladu esenciální matice E_C se aplikuje diagonální matice $S = s, s, 0$. Rozklad bude mít následující tvar [1]:

$$E_C = U \cdot S \cdot V^T \Rightarrow E = U \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot V^T \quad (3.11)$$

Podmínkou použití této osmibodové metody je to, že korespondující si body v prostoru se nesmí nacházet v jedné rovině. Velkou výhodou této metody je, že lze aplikovat i na kamery, které nejsou kalibrované.

3.1.2. Určení matice rotace a translace z esenciální matice

Matice rotace R a translace \hat{t} se získá z esenciální matice E . Matice rotace a translace se rozloží pomocí singulárního rozkladu do následujících tvarů [1]:

$$R = U \cdot (\pm W^T) \cdot V^T \quad (3.12)$$

$$\hat{t} = U \cdot (\pm W) \cdot S \cdot U^T \quad (3.13)$$

kde matice W^T je charakterizována následovně

$$W^T = \begin{bmatrix} 0 & \pm 1 & 0 \\ \mp 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.14)$$

Výpočtem rovnic 3.12 a 3.13 vznikají čtyři různá řešení rovnic. Výběr správného řešení pro R a \hat{t} probíhá na základě rovinné triangulace, kde se triangulovaný bod nachází mezi oběma kamerami. Při nelineární optimalizaci rotace R a translace \hat{t} mohou být takto určené parametry použity jako počáteční hodnoty pro optimalizaci.

Výpočet relativního měřítka

Absolutní transformaci $T_{0:n}$ nelze jednoznačně určit ze sekvenční dvou po sobě jdoucích snímků, nýbrž lze určit pouze relativní měřítko. První možností, jak tohle měřítko určit, je tzv. prostorovou triangulací dvou korespondujících bodů X_k a X_{k-1} , ve dvou po sobě jdoucích snímcích, mezi kterými mohou být určeny relativní vzdálenosti. Relativní měřítko r je charakterizováno následujícím vztahem [1]

$$r = \frac{\|X_{k-1,i} - X_{k-1,j}\|}{\|X_{k,i} - X_{k,j}\|} \quad (3.15)$$

V případě zpřesnění výsledků lze použít relativní měřítko pro jednotlivé korespondující body a průměr všech bodů. Translační vektor \vec{t} je určen pomocí vzdálenostního měřítka.

Shrnutí algoritmu metody 2D do 2D [1]

Algoritmus metody 2D do 2D lze shrnout do těchto následujících bodů:

- 1) Pořízení nového snímku I_k
- 2) Určení korespondujících bodů ze současného snímku I_k a předešlého snímku I_{k-1}
- 3) Výpočet esenciální matice E současného snímku I_k a předešlého snímku I_{k-1}
- 4) Rozklad esenciální matice na rotaci R , translaci \vec{t}_k a formulování pohybu kamery T_k
- 5) Výpočet relativního měřítka r a rozškálování t_k
- 6) Určení celkové grafické transformace $C_k = C_{k-1}T_k$
- 7) Opakuj algoritmus od bodu č. 1

Parametry C_k , C_{k-1} a T_k jsou znázorněny na obrázku č. 3.2.

3.2. Metoda 3D do 3D

U této metody jsou vstupem dvojice snímků. Z nich se následnou prostorovou triangulací získá množina bodů v prostoru. Tyto body jsou k sobě přiřazovány pomocí deskriptorů příznaků a z nich je poté determinován pohyb kamery T_k . U následujícího popisu se předpokládá, že korespondující body byly k sobě správně přiřazeny a chybné hodnoty jsou vyfiltrovány. Při určování parametru T_k se hledá nejmenší vzdálenost mezi dvěma korespondujícími body[1]:

$$\arg \min_{T_k} \sum_i \|\tilde{X}_k^i - T_k \tilde{X}_{k-1}^i\|, \quad k = 1, \dots, N \quad (3.16)$$

Parametr i reprezentuje příznak, hodnoty \tilde{X}_k^i a \tilde{X}_{k-1}^i jsou prostorové souřadnice bodů, kde vektor $\tilde{X} = [x, y, z, 1]^T$. Pro určení translace je nezbytné znát minimálně tři a více korespondujících bodů [18]. Nezbytnou podmínkou pro nalezení správného řešení je, že se společně nesmí nacházet na jedné přímce. Bod \bar{X}_k je těžiště množin bodů, kde jeho souřadnice jsou průměr všech souřadnic bodů množiny. Těžiště z předchozí sekvenční se poté značí \bar{X}_{k-1} .

V první řadě se určí rotační matice R . Z množin bodů a následně i z těžišť se určí matice H , na kterou se aplikuje singulární rozklad matice [14]:

3.3. METODA 3D DO 2D

$$H = \sum_k (X_k - 1 - \bar{X}_{k-1})(X_k - \bar{X}_k) = USV^\top, \quad k = 1, \dots, N \quad (3.17)$$

Z takto získaných hodnot se určí rotační část matice R [1]:

$$R = VU^\top \quad (3.18)$$

Determinant matice R musí být ovšem roven jedné. Pokud determinant matice je roven -1, je výsledek neplatný [14].

Poté se určí translační vektor. Ten je definován rotační maticí R a body, které definují těžiště ze současné a předešlé sekvence [1]:

$$t_k = \bar{X}_k - R\bar{X}_{k-1} \quad (3.19)$$

Pokud nejsou známy nejistoty prostorových bodů, mohou být určeným hodnotám přiřazeny jednotlivé váhy podle článku [19]. Vypočítané transformace mají absolutní měřítko, a tak může být trajektorie vypočítána přímo z transformací.

Shrnutí algoritmu metody 3D do 3D [1]

- 1) Pořízení snímků ze stereo kamery $I_{l,k-1}$, $I_{r,k-1}$, $I_{l,k}$ a $I_{r,k}$
- 2) Výběr a přiřazení korespondujících bodů mezi $I_{l,k-1}$ a $I_{l,k}$
- 3) Prostorová triangulace korespondujících bodů pro každou stereo dvojici
- 4) Výpočet transformace T_k z těžišť X_{k-1} a X_k
- 5) Určení celkové grafické transformace $C_k = C_{k-1}T_k$
- 6) Opakování algoritmu od bodu č. 1

3.3. Metoda 3D do 2D

Vstupem do této metody je dvojice snímků ze stereokamery a trojrozměrné body získané prostorovou triangulací z předchozího snímku. Transformace T_k je spočítána z korespondencí X_{k-1} a průměty p_k v promítací rovině. Z bodů X_{k-1} se určí teoretický průmět p_{k-1} v obraze I_k . Transformace T_k se určí minimalizací rozdílu mezi teoretickým průmětem \hat{p}_{k-1}^i a nalezeným průmětem p_k [1]:

$$\arg \min_T \sum_i \|p_k^i - \hat{p}_{k-1}^i\|^2 \quad (3.20)$$

Tento problém se v literatuře označuje jako PnP (Perspective from n Points). V současnosti se používá spousta metod z nichž nejznámější jsou metody P6P a P3P. U metody P3P se využívá nejmenšího počtu bodů, ze kterého vychází čtyři možná řešení. Zároveň platí podmínka, že body nesmí ležet na jedné přímce. V současnosti je standardní metodou pro robustní určení pohybu.

Při použití šesti a více bodů lze tuto úlohu řešit jako soustavu lineárních rovnic.

3.3.1. Metoda P6P

U této metody se určuje matice $P_k = [R|\vec{t}]$, která je složena z rotační části R a z translační části \vec{t} . Výpočet rovnice je charakterizován následující rovnicí [1]:

$$AP = 0 \quad (3.21)$$

\Downarrow

$$\begin{bmatrix} 0 & 0 & 0 & 0 & -x & -y & -z & -1 & x\tilde{v} & y\tilde{v} & z\tilde{v} & \tilde{v} \\ x & y & z & 1 & 0 & 0 & 0 & 0 & -x\tilde{u} & -y\tilde{u} & -z\tilde{u} & -\tilde{u} \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = 0 \quad (3.22)$$

kde matice A má rozměry $2n \times 12$, body x, y a z jsou prostorové body získané prostorou triangulací předchozího obrazu, \tilde{u} a \tilde{v} jsou homogenní obrazové souřadnice daného objektu. Použitím šesti a vícero bodů, vznikne soustava rovnic. Matici A lze rozložit singulárním rozkladem matice:

$$A = USV^\top \quad (3.23)$$

Poslední sloupec matice V představuje hledaný vektor s prvky pro matici P . Při normalizaci hodnoty vektorů, která se provádí pro zpřesnění výsledku, je doporučeno, aby těžiště bodů ležících v jedné rovině bylo v počátku a průměrná vzdálenost bodů od počátku se pohybovala kolem $\sqrt{2}$. Normalizace závisí na rozdělení bodů v prostoru. Pokud jsou body blízko, doporučuje se umístit střed bodu do středu, aby průměrná souřadnice byla $[1, 1, 1, 1]^\top$. Tyto normalizace jsou definované následujícími rovnicemi [14].

$$\vec{x}_n = N_V \cdot \vec{x} \quad (3.24)$$

$$\vec{X} = N_U \cdot \vec{X} \quad (3.25)$$

Výslednou matici P lze určit následující matematickou rovnicí [14].

$$P = N_V^{-1} \cdot P_n \cdot N_U \quad (3.26)$$

Shrnutí algoritmu 3D do 2D [1]

- 1) Inicializační krok
 - 1.1) Pořízení snímků I_{k-2}, I_{k-1}
 - 1.2) Výběr a přiřazení příznaků ze snímků I_{k-2}, I_{k-1}
 - 1.3) Triangulace příznaků ze snímků I_{k-2}, I_{k-1}
- 2) Krok v každé iteraci
 - 2.1) Pořízení nového snímku I_k
 - 2.2) Výběr a přiřazení příznaků z předchozího snímku I_{k-1}
 - 2.3) Výpočet pozice kamery z korespondujících příznaků pomocí metody PnP
 - 2.4) Triangulace všech nových příznaků mezi snímky I_k a I_{k-1}

4. Detektory a deskriptory příznaků

Při zpracování obrazů a jejich vyhodnocování je nezbytnou součástí detekce a popis příznaků. Zejména při SLAM mapování jsou používány lokální fotometrické příznaky, což jsou význačné body, které se vyznačují rozdílnou intenzitou, barvou nebo texturou vůči okolí. V první řadě se tyto příznaky detekují a poté je pro ně spočítán deskriptor, kterým je příznak jednoznačně identifikován. Pomocí těchto deskriptorů je možné detekovat příznaky z různých pohledů kamery a natočení. Světelné podmínky mají zcela zásadní vliv. Čím víc je změna osvětlení výraznější, tím méně pravděpodobné, že se samé příznaky podaří detekovat. Existuje nespočetná řada metod, kterými lze detekovat a popsat jednotlivé příznaky. Ty se sebe liší vlastními vnitřními parametry, robustností, počtem a charakterem detekovaných příznaků.

Metody mohou detekovat příznaky dvěma způsoby a to na základě detekci rohů (corners) nebo na základě kontrastních objektů (blobs). Detekce hran patří mezi starší metody. Výhodou této metody je invariantnost vůči natočení příznaku a nízká výpočtová náročnost. Nevýhodou je, že roh, který je definován jako průsečík dvou hran, je hůře velikostně invariantní. Příkladem takových detektorů je Harrisův detektor, Shi-Tomasi či FAST detektor. Detekce založené na základě kontrastních objektů detekuje příznaky lišící se svou intenzitou, barvou nebo texturou. Mezi kontrastní detektory patří například SIFT, SURF či CENSURE [1]. Dle článku [20] je doporučeno použít SIFT, který má v sobě implementovaný detektor na základě kontrastních objektů, a BRISK, ve kterém je implementovaný detektor na základě detekci rohů. Na základě tohoto doporučení budou v této analýze použity pouze tyto dvě metody.

4.1. SIFT

Metoda SIFT (Scale-Invariant Feature Transform) je metoda, která obsahuje detektor i deskriptor kontrastních objektů. Tato metoda je navržena tak, aby detekovala příznaky nezávisle na jeho rozměrech. Výhodou je také to, že se tato metoda vyznačuje větší robustností vzhledem k zašumění obrazu, změně osvětlení, různým afinním transformacím a navíc jsou detekované příznaky jednoznačně popsány, tzn. je málo pravděpodobné, že by mohlo dojít k záměně.

Algoritmus SIFT se dělí do těchto čtyřech částí:

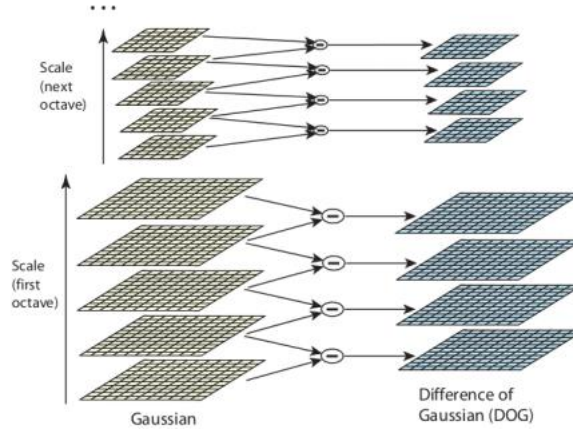
1) DETEKOVÁNÍ JASOVÝCH EXTRÉMŮ NA SNÍMKU:

Detekování extrémů na snímků na základě kontrastních rozdílů je založeno na diferenci Gaussovy funkce. Gaussova funkce je charakterizována následujícím vztahem [21]:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma} \cdot e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (4.1)$$

kde σ je značí rozptyl, a souřadnice x a y představuje pozici. V následujícím kroku se konvolucí bodů I provede určení funkce L , která jako následek rozostření obrazu [21]:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (4.2)$$



Obrázek 4.1: Rozdělení do oktáv [21]

Před výpočtem jsou jednotlivé difference podle obrázku 4.1 rozděleny na oktávy, pro které platí, že následující oktáva má dvojnásobný rozdíl σ . Pro výpočet nalezení extrémů v jedné oktávě je nutné $s + 3$ snímků, kde s představuje počet intervalů, na které je oktáva rozdělena. Konstanta k v rovnici 4.4 je charakterizována následujícím vztahem [21]:

$$k = 2^{\frac{1}{s}} \quad (4.3)$$

Následná difference D je pak charakterizována jako difference funkcí L s různými hodnotami σ [21].

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (4.4)$$

Pomocí těchto diferencí lze zajistit, aby bod byl při různých velikostech detekovatelný.

Následná difference Gaussovy funkce se pak získá z rozdílu po sobě jdoucích snímků určité oktávy. Extrémy jsou určeny v nově získaných snímcích reprezentující další oktávu. Extrém je definován jako bod, který je extrémem za podmínky, že je extrémem vůči okolním osmi pixelům na snímku a osmnácti pixelům, které se nachází v sousedních snímcích.

2) LOKALIZACE PŘÍZNAKŮ DETEKOVANÉ NA ZÁKLADĚ NALEZENÝCH EXTRÉMŮ:

V tomto kroku se definuje umístění a relativní velikost oblasti. Příznaky, které jsou slabé nebo na hranách jsou v tomto kroku vyfiltrovány. Z důvodu zvýšení subpixelové přesnosti je difference aproximována Taylorovým rozvojem druhého řádu, pro kterou je následnou derivací nalezen extrém funkce. Poloha ovšem nesmí být větší než 0.5, jinak je extrém považován za nedostatečně kontrastní a je tedy vyloučen.

Poslední fází tohoto kroku je vyloučení bodů nacházející se na hranách, neboť tyto příznaky nelze považovat za spolehlivé. Tento proces se uskutečňuje pomocí Hessovy matice 4.5, kterou se vypočítá zakřivení plochy v okolí určeného extrému a porovná se s podmínkou 4.6 [21].

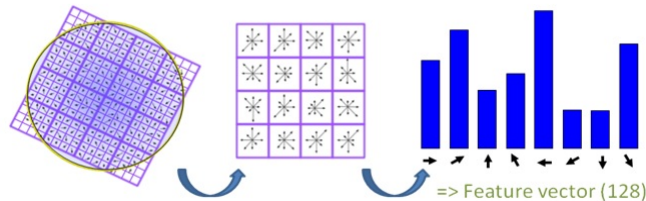
$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix} \quad (4.5)$$

$$\frac{\text{trace}(H)^2}{\det(H)} < \frac{(r + 1)^2}{r} \quad (4.6)$$

3) URČENÍ ORIENTACE K PŘÍZNAKU: Aby byla zajištěna invariance vůči rotaci, jsou v tomto kroku přiřazeny orientace ke každému nalezenému bodu podle směrů lokálních

4.2. BRISK

gradientů. Ty jsou vypočítány pomocí parciálních diferencí, kterými je určena nejen velikost, ale i směr. Úhel je v této metodě rozdělen do třiceti šesti částí po deseti stupních. Z každé této části se určí maximum. Pokud druhá nejvyšší hodnota je vyšší než 80%, jsou pro tento příznak určeny dvě hodnoty. Úhly orientací se určí tak, že se maximum histogramu a jeho dvěma nejbližšími sousedy proloží polynom druhého stupně, kterým se poté určí extrém, tedy přesnější orientace.



Obrázek 4.2: Znázornění deskriptorů pro metodu SIFT [22]

4) DEFINOVÁNÍ DESKRIPTORU PŘÍZNAKU:

Po předchozím kroku je zapotřebí jednoznačně popsat lokální gradienty. Okolí příznaků je transformováno podle vypočtené orientace k tomu, aby příznaky byly invariantní vůči rotaci. V této metodě je okolí 16×16 pixelů rovnoměrně rozděleno do čtyř částí. V jednotlivých bodech se určí gradient váhovan kruhovým Gaussovským oknem. Pro každou část je vypočten histogram gradientů s osmi směry. Hodnoty histogramů jsou parametry deskriptoru s celkem se sto dvaceti osmi hodnotami ($16 \text{ částí} \times 8 \text{ kategorií}$). Aby se omezilo vlivu intenzity osvětlení, jsou vypočítané vektory normovány.

Přiřazování příznaků na základě rozdílu vzdáleností se eliminují špatně přiřazené příznaky. Podle experimentu [23] bylo zjištěno, že pokud poměr minimální vzdálenosti nejbližšího příznaku a dalšího nejbližšího příznaku vůči přiřazovanému příznaku je větší než 0.8, tak dochází k potlačení okolo 90% špatně přiřazených příznaků a přibližně 5% správně přiřazených dvojic.

4.2. BRISK

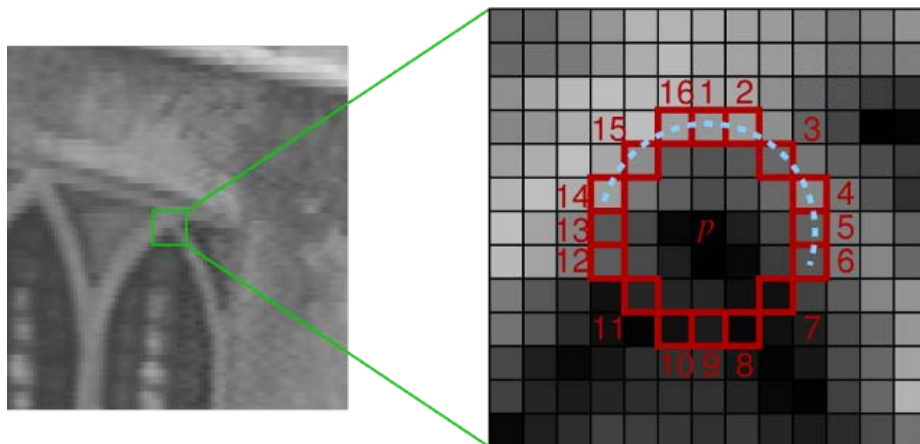
Deskriptor BRISK (Binary Robus Invariant Scalable Keypoints) je binární deskriptor, který detekuje rohy. Využívá metodu FAST (Features from accelerated segment test) [24] a deskriptory porovnává na základě bitových operací. Tato metoda byla vytvořena švýcarskou univerzitou ETH (Eidgenössische Technische Hochschule) v Zürichu.

Porovnávání deskriptorů se uskutečňuje na základě bitových operací. Dle publikace [25] a podle experimentálního zjištění [20] je tato metoda prokazatelně rychlejší než samotná metoda SIFT.

Nalezení příznaku a detekce

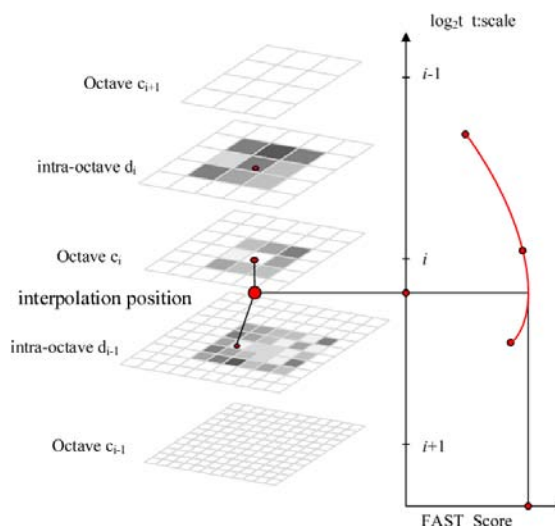
Detekce rohu je stejná jako u metody FAST (Features from accelerated segment test) [24]. Při detekování se pixely prohledávají na kružnici okolo testovaného bodu znázorněného na obrázku 4.3. Daný bod je vyhodnocen jako roh, pokud pro určitý počet po sobě jdoucích bodů v kruhu platí, že intenzita splňuje jednu z podmínek vyjádřenou v rovnici 4.7.

$$I_p + t < I_k, I_p - t > I_k \quad (4.7)$$



Obrázek 4.3: Ukázka detekce rohů na snímku [24].

I_k je intenzita jednotlivých bodů, I_p je intenzita konkrétního bodu a t je zvolený parametr. Měřítkový prostor je tvořen pyramidou obdobně jako u metody SIFT. Zde se také využívá rozdělení do oktáv kvůli omezení vlivu velikosti. V této diskretizaci používá BRISK větší intervaly. Skutečné měřítko se odhaduje tzv. interpolací znázorněnou na obrázku 4.4.



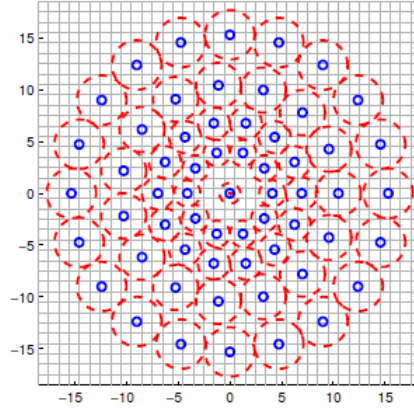
Obrázek 4.4: Znázornění měřítkového prostoru metody BRISK [25]

Deskripce příznaků

Po detekci rohových objektů a jejich následné vytvoření sady příznaků dochází k samotnému popisu. BRISK deskriptor je tvořen binárním řetězcem. V této fázi okolí každého nalezeného příznaku je vzorkováno kruhem o určitém množství bodů N , kde je každé okolí vyhlazeno Gaussovým jádrem z důvodu odstranění šumu. Samotné vzorkování je zobrazeno na obrázku 4.5.

Bitový vektor deskriptoru v dané oblasti d_k je tvořen párem vzorkovacích bodů (p_i, p_j) , která se nacházejí v již v předem definované minimální oblasti. Právě z těchto porovnání je sestaven vektor deskriptoru d_k . Bit je charakterizován podle následující rovnice [25]:

4.2. BRISK



Obrázek 4.5: Vzorkování okolo keypointu v metodě BRISK [25]

$$b = \begin{cases} 1, & I(p_j, \sigma_j) > I(p_i, \sigma_i) \\ 0 \end{cases} \quad (4.8)$$

Funkce $I(p, \sigma)$ je intenzita vzorkovacího bodu upravena Gaussovským vyhlazovacím jádrem velikosti σ .

Při porovnávání deskriptorů se určují tzv. Hammingovy vzdálenosti, kterou se určí rozdílný počet bitů, určující rozdílnost dvou deskriptorů. K tomuto porovnání postačí jednoduchý logický operátor XOR, kterými se porovnají dva deskriptory a poté se spočítají bity.

5. Přiřazování a filtrování příznaků

5.1. Přiřazování příznaků

Při detekci pohybu je nezbytné přiřazování bodů podle podobnosti deskriptorů. Základem této metody je přiřazení nejbližšího příznaku z předchozích snímků. Při tomto zpracování může docházet k celé řadě chyb, kdy může dojít ke špatnému přiřazení příznaků či se může stát, že jeden příznak může být přiřazen k několika dalším jiným příznakům. Tento způsob přiřazování je vylepšen tím, že přiřazování sobě si nejbližších příznaků probíhá i zpětným způsobem, který se musí shodovat. Tímto způsobem dochází i k vyfiltrování špatných příznaků a k zásadnímu omezení špatných přiřazení. Nevýhodou tohoto způsobu je výkonová a výpočetní náročnost.

5.1.1. Knihovna FLANN

Výpočetní metody na přiřazování příznaků jsou výpočtově náročné, proto je zapotřebí použití jiných metod. Existuje nespočet různých metod. Jednou z dosud nejpoužívanějších metod je seřazení příznaků do stromové struktury, která umožňuje rychlejší prohledávání. V programu RTAB-Map je implementována knihovna FLANN (Fast Library for Approximate Nearest Neighbours). V případě použití příznaku reálných vektorů SIFT, je v případě této analýzy v této diplomové práci použit KdTree. V případě použití binárního deskriptoru BRISK se využívá metoda hierarchical clustering tree.

Metoda KdTree

Metoda KdTree je upravená metoda k-d stromu (k-dimensional), kde dimenze, podle které se dělí stromová struktura, je vybrána zcela náhodně z celkového počtu dimenzí. V této metodě se vytvoří několik těchto stromů. V případě, kdy dochází k vyhledávání těchto uzlů, jsou tyto stromy paralelně prohledávány. Ze všech uzlů je poté vybrán ten, který je nejbližší k listu. Konečný uzel je vybrán až poté, kdy je nalezeno vícero těchto listů.

Metoda Hierarchical Clustering Tree

V této metodě jsou uzly tvořeny zcela náhodně ze shluků k . Ostatní body jsou rozděleny do shluků na základě vzdáleností. Každý strom je taktéž prohledáván zvlášť a v případě, že se na výstupu nachází dva a více bodů, je pro každý strom vytvořena priorita. Výstupem jsou přiřazené příznaky s nejmenší vzdáleností.

5.2. Filtrování příznaků

5.2.1. RANSAC

Snahou při přiřazování příznaků je vyřadit chybná přiřazení. Ta mohou ovlivnit vyhodnocování trajektorie pohybu daného objektu. Doposud nejpoužívanější metodou pro vyfiltrování těchto chybných přiřazení příznaků v oblasti vizuální odometrie je metoda RANSAC. Na rozdíl od aproximací, které jsou ovlivnitelné chybnými příznaky, dokáže RANSAC pracovat i s velkým množstvím špatně přiřazených hodnot. Jelikož se v této metodě používá náhodný výběr bodů, nelze nikdy plně determinovat její chování, což je její hlavní nevýhodou.

V této metodě se pro náhodně vybrané množiny detekovaných bodů určí pohyb daného objektu. Za správný výsledek je považován ten s nejlepším ohodnocením. Metoda RANSAC je charakterizována následující rovnicí [14]:

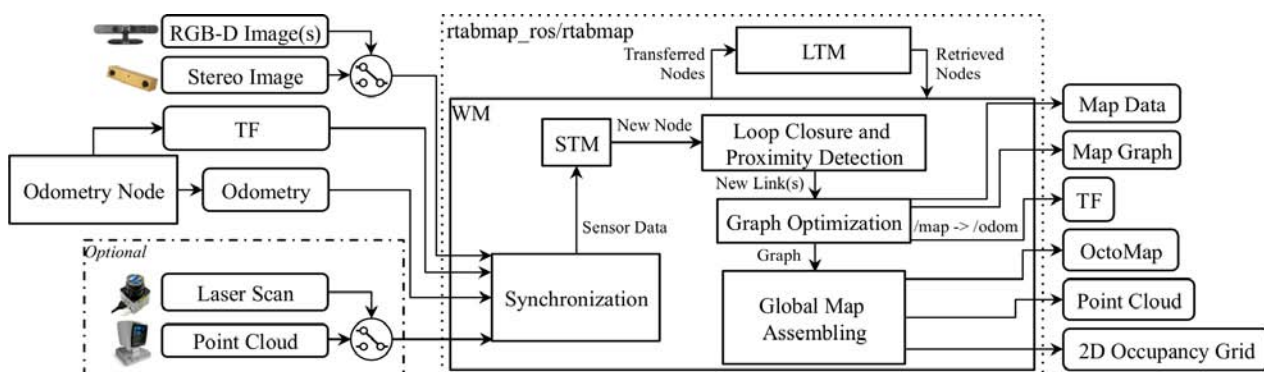
$$N = \frac{\log(1 - p)}{\log(1 - (1 - e)^s)} \quad (5.1)$$

kde p je pravděpodobnost koeficientu úspěchu, s je počet bodů nezbytný pro definování modelu pohybu a e představuje poměr špatně přiřazených bodů vůči celé množině bodů. Z rovnice je zřejmé, že nezbytný počet kroků N exponenciálně narůstá s nezbytným množstvím bodů pro určení pohybu objektu. Jelikož s rostoucím počtem bodů roste i výpočtová náročnost, doporučuje se množství bodů snížit a to eliminováním šumu a případně omezením počtu stupňů volnosti.

6. RTAB-Map

RTAB-Map (Real-Time Appearance-Based Mapping) [26], [27], [28], [29], [30], [31] je otevřená zdrojová knihovna pro simultánní lokalizaci a mapování v systému ROS [32], ve které je implementována detekce uzavřených smyček. Výstupem je graf, který obsahuje tzv. uzly s odkazy na údaje (např. o nalezených příznacích, o poloze), které jsou uloženy v databázi. RTAB-Map podporuje širokou škálu druhu mapování skrze RGB-D, stereokameru, Lidar a jejich různých kombinací.

Obecná funkce RTAB-Map je znázorněna v následujícím blokovém diagramu.



Obrázek 6.1: Blokový diagram RTAB-Map v systému ROS [26]

Do RTAB-Map vstupují data z topiců (témat), na jejichž základě probíhají veškeré komunikace v systému ROS. Každý topic představuje přístupný komunikační bod mezi uzly, kde se nacházejí zprávy určitého datového typu, který daný topic přijímá nebo odesílá. V topicu „TF“, se nacházejí data, ve kterých jsou zadefinovány veškeré pozice v souřadnicovém systému. Dalším vstupem jsou odometrie se třemi nebo se šesti stupni volnosti a obrazová data ze stereokamery nebo hloubkové RGB snímky. Do programu mohou zároveň vstupovat data z laserového scanu, z 2D lidarů nebo data z point cloudu z 3D lidarů. Ta vstupují do pracovní paměti (WM - Working Memory), kde dojde k jejich synchronizaci. Poté se zpracují v STM (Short-Time Memory). Po zpracování dat dochází v grafu k vytvoření nového uzlu. Poté se aktivuje detekce uzavřených smyček, kterou algoritmus zjišťuje, zda se sledovaný objekt v dané poloze nenacházel. Tato lokalizace probíhá i při mapování, aby nedocházelo k duplicitě uzlů. Po tomto kroku dojde k optimalizaci a k úpravě samotného grafu. Poté se data z takto upraveného grafu použijí do sestavení globální mapy. V případě mapování se data uloží do LTM (Long-Time Memory), což je paměť, ve které se nachází data z jednotlivých uzlů a ze které se čerpá pro lokalizaci uzavřených smyček. Výstupem jsou Map Data obsahující mapová data z posledně vzniklého uzlu popřípadě i data ze senzoru, dále korekce odometrie (TF), OctoMap a PointCloud.

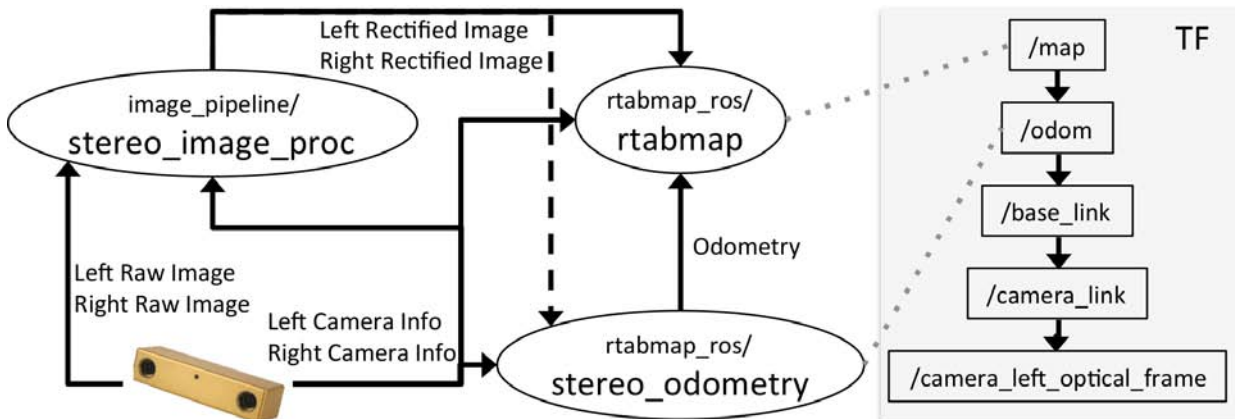
Pokud do RTAB-Mapu nevstupují žádná data z odometrie, je možné použít vizuální odometrii či odometrii pro lidar implementovanou v knihovně RTAB-Map. Při analýze stereokamery ZED jsou použita výstupní data z vizuální odometrie ZEDfu vyvinutá přímo výrobcem Stereolabs.

6.1. Synchronizace dat

RTAB-Map je schopen zpracovávat data z několika vstupních topiců (hloubkové RGB snímky, snímky ze stereokamery, data z odometrie, data ze 2D laser scanu, vstupní 3D

6.2. STM

point cloud a vstupní uživatelská data). Nezbytně požadovanými vstupy pro správnou funkčnost SLAM mapování v RTAB-Map knihovně jsou snímky buď z jedné či více RGB-D kamer nebo ze zkalibrované stereokamery. Společně se vstupními obrazovými snímky jsou také nezbytná přesná vstupní data o pořízených snímcích, ve kterých jsou nezbytné informace pro zpracování snímků (čas, kalibrační parametry apod.). Nezbytné je také řazení souřadnicových systému. Na následujícím obrázku je znázorněn blokový diagram SLAM v RTAB-Map při použití stereokamery.



Obrázek 6.2: Blokový diagram SLAM v RTAB-Map při použití stereokamery [26]

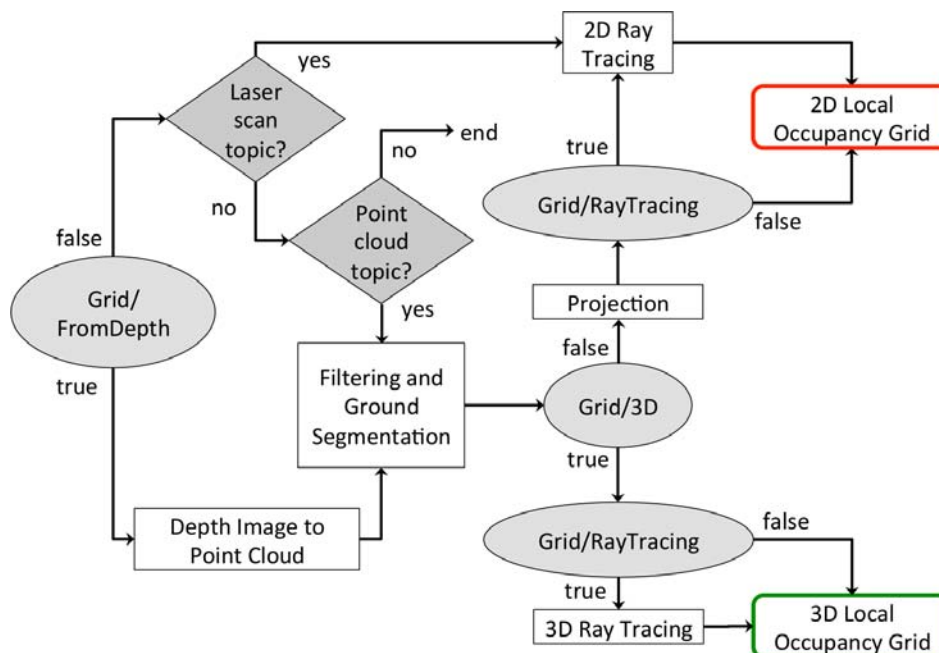
Na obrázku 6.2 jsou znázorněny výstupy ze stereokamery. Snímky a data ze stereokamery vstupují do procesu, kde se snímky rektifikují a vstupují do RTAB-Map, kde jsou následně společně s daty zpracovány. Data ze stereokamery slouží kromě k rektifikaci také ke zpracování odometrie, které následně vstupují do RTAB-Mapu. Napravo od blokového diagramu je znázorněna konfigurace jednotlivých sensorů. Tečkované čáry značí, kde jsou daná data opublikována.

Nikdy nelze plně zaručit, aby data ze sensorů byla pořízena ve stejný čas. V RTAB-Mapu existují dva typy synchronizace a to „exact“ a „approximate“. Tzv. exaktní synchronizace vyžaduje, aby data ze stejného senzoru byla pořízena ve stejný časový okamžik. Aproximační synchronizace je použita v tom případě, kdy se data pořizují z jiných sensorů. Tento typ synchronizace porovnává, v jakém okamžiku byla daná data pořízena, a zesynchronizuje je podle toho, zda se časové zpoždění nachází v toleranci.

U stereokamery ZED není z důvodu obchodního tajemství dostupná žádná dokumentace, ve které jsou popsány veškeré metody a procesy použité pro zpracování dat ve vizuální odometrii ZEDfu. Z tohoto důvodu jsou v této kapitole popsány procesy a metody, použité při analýze SLAM v RTAB-Map.

6.2. STM

Short-Time memory (STM) je použita pouze při vytvoření nového uzlu. Vytvoření lokální mapy je popsáno na následujícím diagramu 6.3.



Obrázek 6.3: Blokový diagram SLAM v RTAB-Map při vytváření lokální mapy [26]

V případě použití stereokamery je point cloud vytvořen tzv. „Block Matching“ algoritmem [33]. Velikost lokální hloubkové mapy je určena použitým senzorem. Dle diagramu je zřejmé, že musí být definováno, zda vygenerovat 3D či 2D hloubkovou mapu. Ze 3D hloubkové mapy je možné vygenerovat i 2D mapu, ale nevýhodou je její paměťová náročnost. Ze 2D mapy nelze zpětně vygenerovat 3D mapu, ale oproti 3D mapě není paměťově náročná. V blokovém diagramu 6.3 se nacházejí tyto funkce:

2D Ray Tracing Vytváření mapy pomocí laserového snímání okolí.

Depth Image to Point Cloud V této funkci dochází převodu bodů, nalezené pomocí stereokamery, do 3D point cloudu.

Filtering and Ground Segmentation V tomto procesu dochází k převzorkování point cloudu pomocí „Voxel Grid“ filtrace [34]. Velikost voxel filtrace je fixována na velikost gridové velikosti buněk. Základní rovina se rozsegmentuje z point cloudu, poté jsou ke každé buňce vypočítány normály. Pokud úhel normály vůči ose z přesáhne fixní úhel, jsou zarovnány k vůči zemi základní rovině. Pokud tento úhel přesáhne, jsou body brány jako překážky.

Projection Pokud je vypnutá funkce GRID/3D, jsou veškeré detaily znázorněny do rovinné mapy. Opět se aplikuje „voxel grid“ filter ke spojení nalezených bodů a vyplnění prázdných prostor v mapě mezi robotem a nalezenými překážkami. Pokud není použita funkce „2D Ray Tracing“ a pokud nejsou rozsegmentovány žádné body v point cloudu, nedojde k žádnému vykreslování překážek.

6.3. DETEKCE UZAVŘENÝCH SMYČEK A PŘÍBLIŽNÁ LOKALIZACE

3D Ray Tracing OctoMapa je vytvořena ze samostatných lokálních gridů v rozsahu robota. OctoMapa provádí tuto funkci a detekuje prázdné buňky mezi kamerou a sledovaným objektem. OctoMapa je převedena zpět do lokální gridovaného formátu s prázdnými buňkami či s buňkami základní roviny nebo překážek.

6.3. Detekce uzavřených smyček a přibližná lokalizace

Detekce uzavřené smyčky je založena na metodě „bag-of-words“ popsaná v článku [35]. V případě, že se vytvoří nový uzel, STM vyextrahuje příznaky a uloží je do databáze. V případě použití vizuální odometrie F2F (Frame to Frame) nebo F2M (Frame to Map) implementované v RTAB-Map lze použít příznaky přímo z vizuální odometrie. Dle článku [26] v případě detekce uzavřených smyček není třeba velkého množství příznaků jako u vizuální odometrie. Před uložením příznaků do databáze se porovnají získané příznaky s předešlými příznaky uložené v databázi. Ovšem může také dojít k tomu, že získané příznaky mohou být podobné s předešlými. Z tohoto důvodu je nastavena velikost STM, kterou se určuje, kolik dat se musí nasbírat než dojde k uložení příznaků do databáze. K výpočtu pravděpodobnosti, zda se vytvořený uzel nachází mezi uloženými příznaky, je pomocí Tf-IDF metodice vypočítaná četnost a důležitost, který použita pro aktualizace Bayesova filtru. Pokud se pomocí Bayesova filtru potvrdí hypotéza, že se v mapě daný objekt již nacházel, je smyčka detekována a spočítána aktuální poloha. Pokud se daný uzel nepodaří detekovat, je v případě módu mapování vytvořena v mapě a poté uložena do databáze.

Přiřazení k uzlu v blízkosti aktuální pozice daného objektu je popsáno v článku [36]. S tímto typem lokalizace je možné detekovat lokalizaci v případě, že se objekt nachází nejen v tom samém, ale v odlišném směru, čehož samotná detekce uzavřených smyček není schopná detekovat. Zatímco je detekce uzavřených smyček závislá na velikosti paměti, celkový systém přibližné lokalizace je vázán na uzel, který je k robotovi v již vytvořeném grafu nejbližší. Uživatelem může být nastaven počet sledovaných uzlů v aktuální blízkosti. Pokud se robot lokalizuje na jiné poloze a do této polohy se také přemístí, lze použitím tohoto „thresholdu“ zamezit tomu, aby používal uzly z oblasti, kde se nacházel před lokalizací.

V RTAB-Mapu existují dvě fáze SLAM. První z nich je mapování a lokalizace, kde danou oblast mapuje a dochází k samotnému ukládání dat do databáze. Lokalizace je aktivní z toho důvodu, aby nedošlo k ukládání duplicitních uzlů. Druhou fází je samotná lokalizace, kdy nedochází k žádnému vytváření uzlů a ani k ukládání dat, ale pouze k samotné lokalizaci.

6.4. Optimalizace grafu

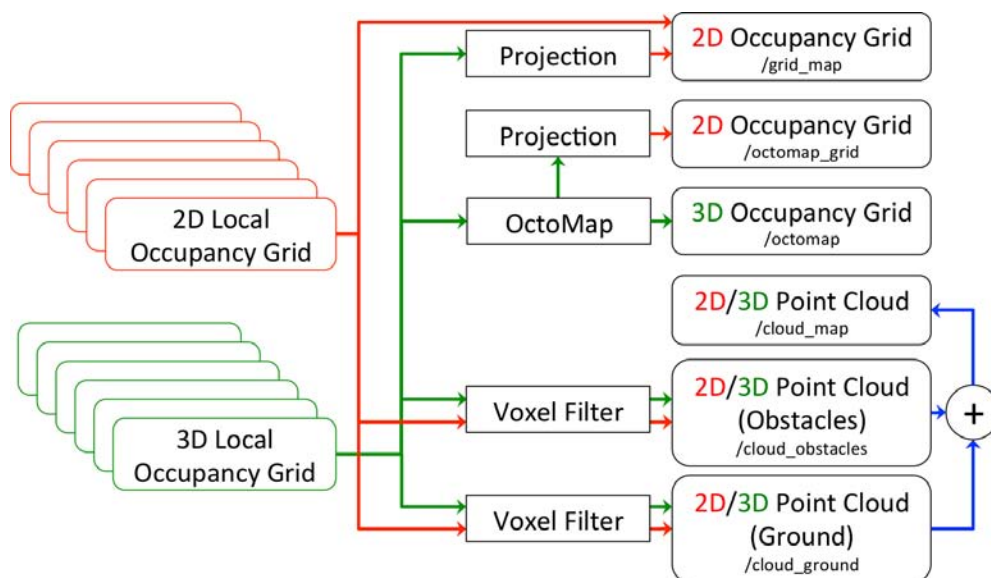
Pokud dojde k detekci u uzavřené smyčky či k přibližné lokalizaci, dochází k optimalizaci grafu, aby se eliminovaly chyby. V RTAB-Map jsou implementovány tři typy grafových optimalizací: TORO [37], G²O [38] a GTSAM [39]. G²O a GTSAM konvergují rychleji než TORO, ale zato jsou méně robustní v případě, že se stejná mapa vytváří vícekrát než jednou. TORO je naopak méně citlivý vůči kovarianci odometrie. V článku [26] je doporučeno používat G²O a GTSAM použít pro samostatnou mapu, zvláště když se jedná

o mapu se šesti stupni volnosti. V případě, že se dojde k vícero mapování stejné oblasti, je GTSAM více robustní než G^2O .

I přes veškeré zabudované funkce může dojít ke špatné detekci uzavřených smyček. Aby se tomuto zamezilo, je v RTAB-Map zabudovaný nový parametr, u kterého se nastavuje maximální optimalizační chyba. Pokud transformace uzlů po optimalizaci přesáhne tuto chybu nebo je translační kovariance příliš velká, veškeré detekované uzavřené a přibližná určení polohy způsobenou detekcí nového uzlu jsou vymazány.

6.5. Vytváření globální mapy

Vytváření mapy je sestavováno ukládáním dat získaných z lokální mapy. Je možné vytvářet 3D mapu a 2D mapu. Výhodou 3D mapy je to, že je z ní možné generovat všechny možné typy map. Při vytvoření nového uzlu je lokální mapa transformována do globální mapy, kam se přidávají či popřípadě odstraňují překážky. Pokud je detekována uzavřená smyčka, může být globální mapa korigována podle nově optimalizovaného uzlu. Tento proces je požadován v případě, že daná překážka bude nesprávně odstraněna. Uzavřenou smyčkou může být opět obnovena. Výstupní formát point cloudu je ve formátu využitelný pro širokou škálu jiných programů. Na obrázku je znázorněno blokové schéma, ve kterém je znázorněno vytváření globální mapy.



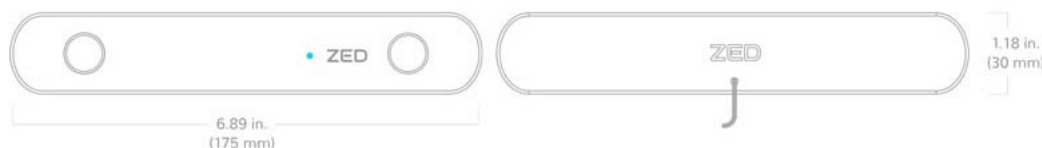
Obrázek 6.4: Blokový diagram SLAM v RTAB-Map při vytváření globální mapy [26]

7. Návrh analýzy a vyhodnocení dat

Stereokamera ZED je vyvíjena společností Stereolabs. Je široce aplikovatelná v dronech, robotech, při různých analýzách pohybu. Také se používá pro virtuální realitu a pro AR. AR (Augment reality) znamená, že přidává digitální elementy do kamerou snímaných objektů. Hlavní funkcí je hloubkové snímání od 0.5 do 20 metrů ve vnitřním či venkovním prostředí, určování aktuální pozice a orientace ve 3D prostoru a mapování prostředí.

Technické parametry udávané výrobcem

ZED Kamera



Obrázek 7.1: Rozměry stereokamery ZED [40]

•**funkce** vysoké rozlišení a vysoká snímkovací frekvence 3D snímání videa hloubkové snímání ve vnitřním i venkovním prostředí do 20 m
snímání pohybu se šesti stupni volnosti
prostorové mapování

•**operační systém:** Windows 7, 8 a 10; Ubuntu 16.04 a 18.04; OS X

•**podporované platformy:** ROS, ROS2, OpenCV, MATLAB, TensorFlow

•**třída:** UVC (Universal Video Class)

	rozlišení	doporučené snímkovací frekvence	rozlišení
• video formát	2.2K	15	4416 × 1242
	1080p	15, 30	3840 × 1080
	720p	15, 30, 60	2560 × 720
	WVGA	15, 30, 60, 100	1344 × 376

•**hloubka snímání** 0.5 – 20 m

•**hloubkový formát** 32-bits

•**přesnost snímání pohybu** ± 1 mm v posunutí a 0.1° v natočení při FPS= 100

•**maximální snímkovací frekvence** 100 Hz

•**metody snímání pohybu** vizuální odometrie nebo SLAM

7. NÁVRH ANALÝZY A VYHODNOCENÍ DAT

•parametry čoček	širokoúhlé s maximálním zorným polem $90^\circ(\text{H}) \times 60^\circ(\text{V}) \times 110^\circ(\text{D})$ a aperturou čočky $f/2.0$
•vzdálenost čoček	120 mm
<hr/>	
•výstupní rozlišení	4M pixely s velikostí pixelu $2 \mu\text{m}$
•velikost senzoru	1/3" CMOS snímač s vysokou fotocitlivostí
•výstupní rozměr	16 : 9 s nastavitelnou výškou snímku
•synchronizace	elektronická synchronizace „Rolling Shutter“
•ovládací možnosti	rozlišení nastavení snímkovací frekvence světla kontrast jas saturace
<hr/>	
•připojení	USB 3.0 s délkou kabelu 1.5 m
•napájení	přes USB 5 V/380 mA

SDK (Software Development Kit)

•procesor:	Dual-core 2,3 GHz nebo výkonější
•RAM:	≥ 4 GB
•grafická karta	Nvidia GPU s vlastností (compute capability) > 3.0

Při analýze byl použit dostupný balíček „zed-ros-wrapper“, ve kterém jsou přístupné všechny funkce z SDK (verze 2.7) pro systém ROS. Před samotnou instalací tohoto balíčku je třeba nainstalovat CUDA Toolkit společně se ZED SDK. V balíčku se nachází řada funkcí pro ZED kameru včetně nahrávání videa do SVO souboru pro pozdější offline analýzu, diagnostikou ZED kamery v případě poruch, hloubkové snímání, vizuální odometrie a různými ukázkovými příklady např. práce v RTAB-Mapu či použití ZED kamery jako laser scanu.

7.1. Návrh analýzy dat

Ze zpracovaných rešerší plyne, že vizuální odometrie a SLAM nefunguje na principu detekce bodů v hloubkových mapách, jak je uvedeno v zadání této diplomové práce, nýbrž pracuje s detekovanými příznaky na snímcích. Proto se v analýze testuje vliv veličin podstatných pro detekci příznaků, na základě kterých se vypočítává odometrie a určuje lokalizace.

7.2. MAPOVÁNÍ

Celá analýza byla provedena v prostředí RTAB-Map, kam kromě snímků a dat z kamery vstupuje také vizuální odometrie ZEDfu vyvíjena samotným výrobcem. Kalibrační parametry jsou dostupné na oficiálních stránkách firmy Stereolabs. V případě SLAM mapování na testování mapování a lokalizace byla využita příznaková metoda z 3D do 2D popsána v kapitole 3.3. Testování probíhalo při použití dvou příznaků a to SIFT a BRISK. K vyhledávání příznaků v databázi bylo použito stromového prohledávání popsané v kapitole 5.1.1. V případě použití detektoru a deskriptoru SIFT byla použita metoda KdTree a v případě použití detektoru a deskriptoru BRISK byla použita metoda Hierarchical Clustering Tree. Obě metody se nachází v knihovně dostupné v OpenCV [41]. Pro optimalizaci grafu je využita metoda G^2O .

Offline zpracování dat probíhalo na počítači se čtyř jádrovým procesorem Intel® Core™i7-7700HQ, s operační pamětí typu DDR4, s frekvencí 2133 MHz a s operační pamětí o velikosti 8 GB. Jako grafická karta byla použita NVIDIA® GeForce® GTX 1050Ti s typem paměti GDDR5 a s její velikostí 4 GB. Celé testování a analyzování dat probíhalo v operačním systému Ubuntu 16.04 s nainstalovaným ROS Kinetic.

7.2. Mapování

Testování probíhalo na trase o celkové délce 250 m za různých podmínek prostředí. Data byla nahrávána pomocí ZED SDK do SVO formátu při rolišení 720p a snímkovací frekvenci 60 Hz. Výběr byl uskutečněn na základě publikace [42], kde podle autorů dochází k přijatelným chybám při hloubkovém snímání. Kromě dat získání dat ze ZEDfu vizuální odometrie a SLAM mapování z RTAB-mapu, byla pomocí funkce „roscat“ sbírána data z kolové odometrie a z GPS.

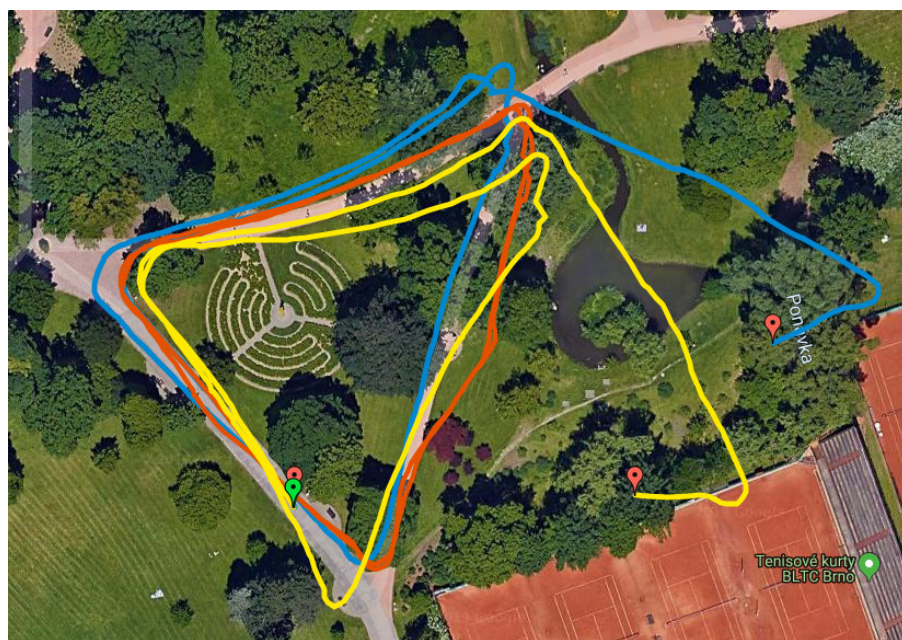
Před spuštěním se kontrolovalo, zda se poloha ZED kamery nacházela v kolmé poloze a před samotným měřením se zjišťoval azimut a počáteční GPS souřadnice. Jelikož u mapování docházelo k vytváření lokální mapy robota v kartézských souřadnicích, kde se souřadnicový systém lišil nejenom počátkem, ale i natočením, byl před každým začátkem měření dat zjištěn azimut a počáteční souřadnice GPS, která byla převedena do UTM systému (Universal Transverse Mercator coordinate system), což je Mercatův systém také v kartézských souřadnicích. Při transformaci lokální mapy robota, byl jeho souřadný systém natočen o naměřený azimut okolo počátku a poté přetransformován podle počátečního bodu do UTM. Následně byly takto transformované body převedeny do GPS souřadnic. Pomocí nich bylo možné znázornit trajektorii na mapě a porovnat tak zmapovanou trajektorii robota s reálnými daty. Při testování byl nebyl omezen maximální počet detekovaných příznaků ani množství porovnávaných příznaků. Frekvence detekování byla nastavena na maximum. Z důvodu vyšší výpočetní náročnosti byl vypnut i voxel filtr. Všechna ostatní nastavení byla ponechána podle defaultního nastavení.

V systému se srovnávalo mapování prostředí za různých podmínek (špatné osvětlení, rychlé natočení kamery, vibrace) a za použití dvou detektorů a deskriptorů SIFT a BRISK. Výsledky byly porovnány se samotnou kolovou odometrií, s naměřenými GPS souřadnicemi, s vizuální odometrií ZEDfu i se SLAM mapováním.

Mapování dne 22. 11. 2018



Obrázek 7.2: Mapování ze dne 22. 11. 2018. GPS data (světle modrá), kolová odometrie (světle fialová).



Obrázek 7.3: Mapování ze dne 22. 11. 2018. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).

7.2. MAPOVÁNÍ

Mapování dne 21. 2. 2019

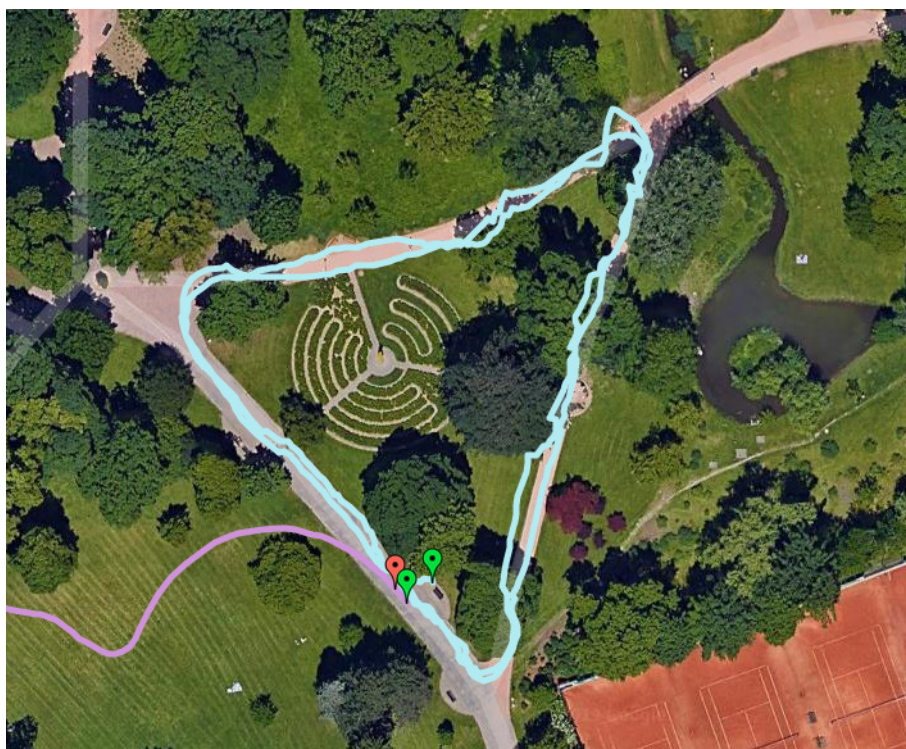


Obrázek 7.4: Mapování ze dne 21. 2. 2019. GPS data (světle modrá), kolová odometrie (světle fialová).



Obrázek 7.5: Mapování ze dne 21. 2. 2019. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).

Mapování dne 7. 3. 2019



Obrázek 7.6: Mapování ze dne 7. 3. 2019. GPS data (světle modrá), kolová odometrie (světle fialová).



Obrázek 7.7: Mapování ze dne 7. 3. 2019. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).

7.2. MAPOVÁNÍ

Mapování dne 21. 3. 2019



Obrázek 7.8: Mapování ze dne 21. 3. 2019. GPS data (světle modrá), kolová odometrie (světle fialová).



Obrázek 7.9: Mapování ze dne 21. 3. 2019. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).

Mapování dne 4. 4. 2019 (dopoledne)



Obrázek 7.10: Mapování ze dne 4. 4. 2019 v dopoledních hodinách. GPS data (světle modrá), kolová odometrie (světle fialová).



Obrázek 7.11: Mapování ze dne 4. 4. 2019 v dopoledních hodinách. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).

7.2. MAPOVÁNÍ

Mapování dne 4. 4. 2019 (odpoledne)



Obrázek 7.12: Mapování ze dne 4. 4. 2019 v odpoledních hodinách. GPS data (světle modrá), kolová odometrie (světle fialová).



Obrázek 7.13: Mapování ze dne 4. 4. 2019 v odpoledních hodinách. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).

Mapování dne 11. 4. 2019



Obrázek 7.14: Mapování ze dne 11. 4. 2019 v dopoledních hodinách. GPS data (světle modrá), kolová odometrie (světle fialová).



Obrázek 7.15: Mapování ze dne 11. 4. 2019 v dopoledních hodinách. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).

7.2. MAPOVÁNÍ

Z analyzovaných dat vyplývá, že mapování trajektorie samotnou vizuální odometrií ZEDfu za použití uzavřených smyček v SLAM funguje velmi spolehlivě za každých měřených podmínek. Při zhoršených světelných podmínkách a to při měřeních konaných 22.11.2018 a 21.2.2019 nejlépe zmapoval trajektorii SLAM za použití detektoru a deskriptoru BRISK. Při ostatních měření byly trajektorie zmapované SLAM v RTAB-Map takřka shodné nezávisle na zvoleném detektoru a deskriptoru. Měření pomocí GPS ani v jedné analýze neodpovídalo zcela přesně trajektorii a v každém měření byla data velmi zatížena šumem. Jednoznačně nejhorší měření trajektorie bylo pomocí kolové odometrie.

7.3. Lokalizace

Lokalizace se testovala na těch mapách, které vyšly ve shodě s trasou. Samotné testování lokalizace probíhalo zvláště pro detektor a deskriptor SIFT a BRISK. Při tomto testu se používala data z různých měření, které bylo prováděno za různých podmínek. Na mapě vlevo jsou vyznačena místa, na kterých došlo k lokalizaci z dat zaznamenaná při dalších měření trasy, která byla provedena hned po mapování. Na mapě vpravo jsou znázorněna místa, došlo k lokalizaci z dat, která se používala při generování ostatních map kromě těch, ze kterých už byla mapa vytvořena. Legenda je znázorněna na obrázku 7.16 a na dalších snímcích 7.17 je znázorněno, za kterých podmínek byla měření provedena.

- data z měření ze dne 22. 11. 2018
- data z měření ze dne 21. 2. 2019
- data z měření ze dne 7. 3. 2019
- data z měření ze dne 21. 3. 2019 (dopoledne)
- data z měření ze dne 21. 3. 2019 (odpoledne)
- data z měření ze dne 4. 4. 2019 (dopoledne)
- data z měření ze dne 4. 4. 2019 (odpoledne)
- data z měření ze dne 11. 4. 2019 (dopoledne)
- data z měření ze dne 11. 4. 2019 (odpoledne)

Obrázek 7.16: Legenda pro grafy lokalizace.



(a) měření 22. 11. 2018



(b) měření 21. 2. 2019



(c) měření 7. 3. 2019



(d) měření 21. 3. 2019 (dopoledne)



(e) měření 21. 3. 2019 (odpoledne)



(f) měření 4. 4. 2019 (dopoledne)



(g) měření 4. 4. 2019 (odpoledne)



(h) měření 11. 4. 2019 (dopoledne)

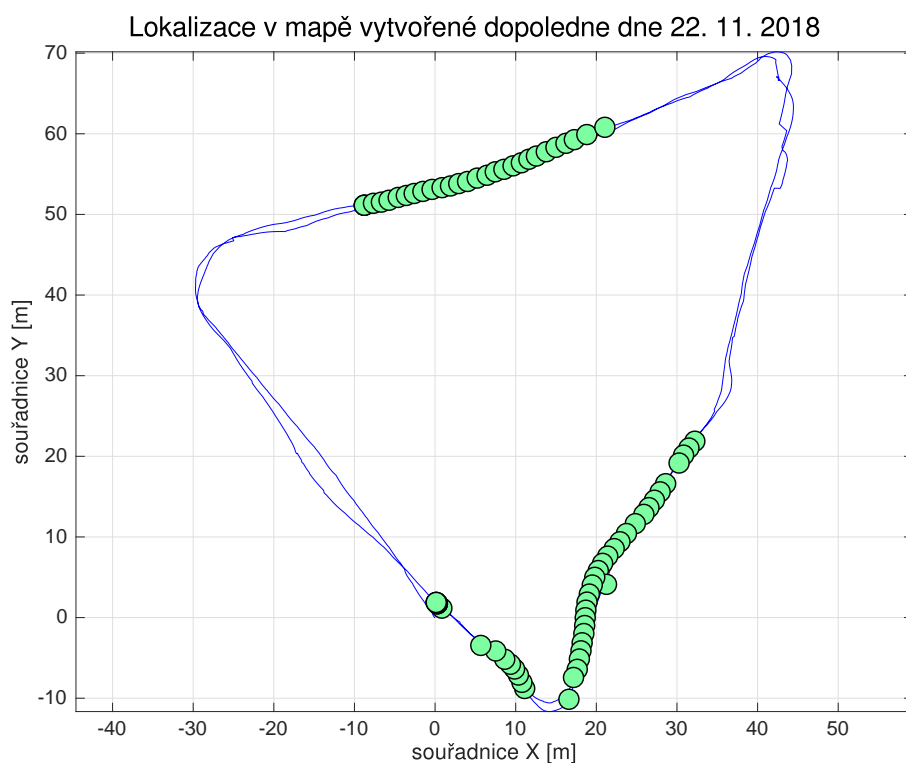


(i) měření 11. 4. 2019 (odpoledne)

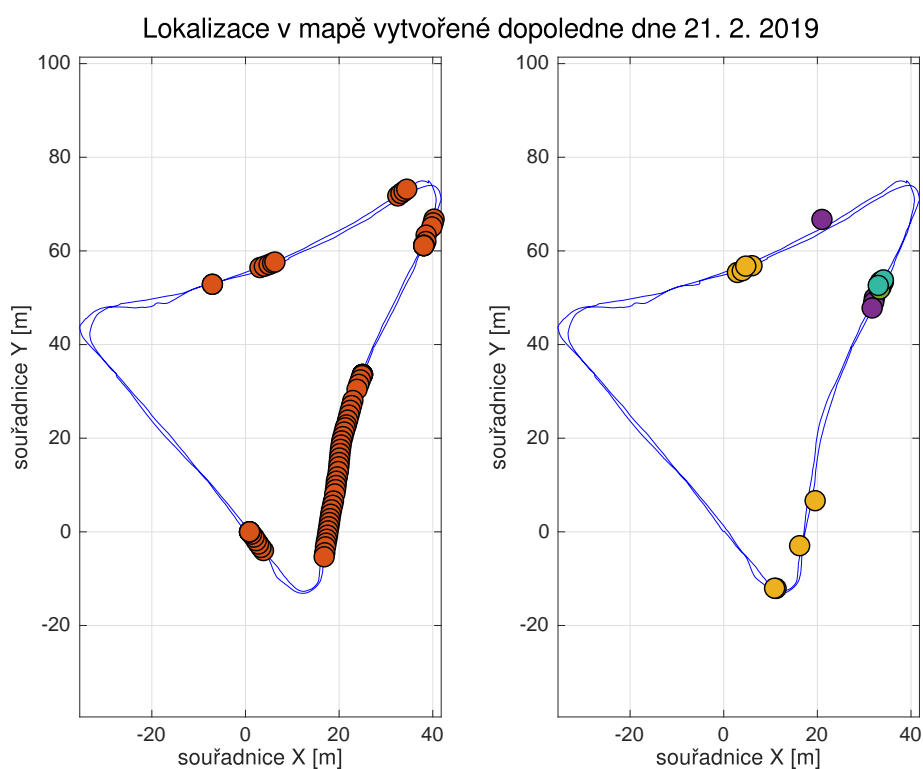
Obrázek 7.17: Ukázky z daných měření

7.3. LOKALIZACE

7.3.1. Lokalizace za použití metody BRISK

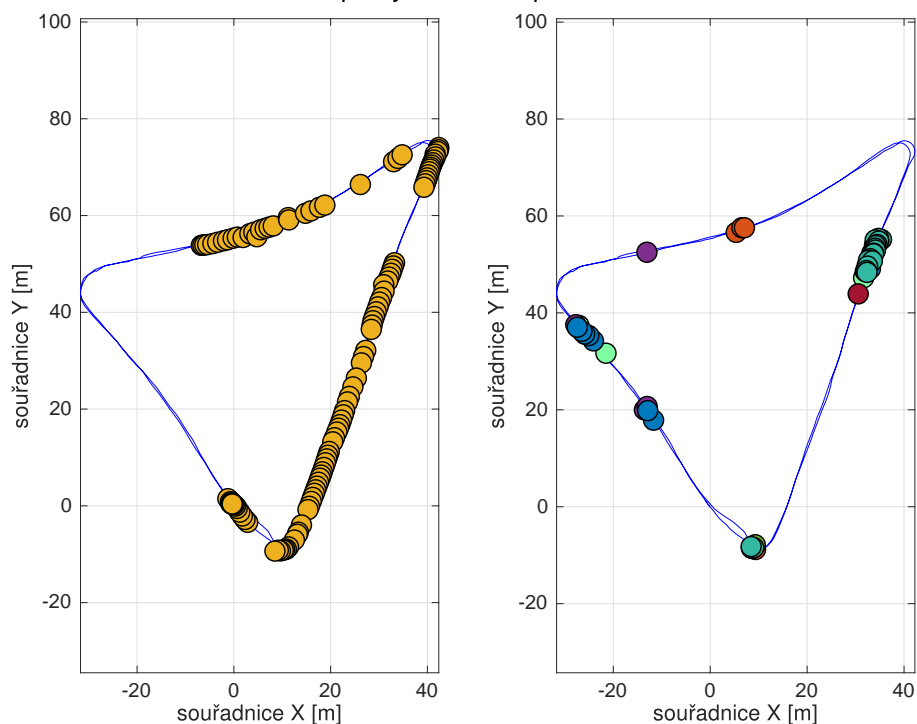


Obrázek 7.18: Lokalizace na mapě vytvořené 22. 11. 2018.



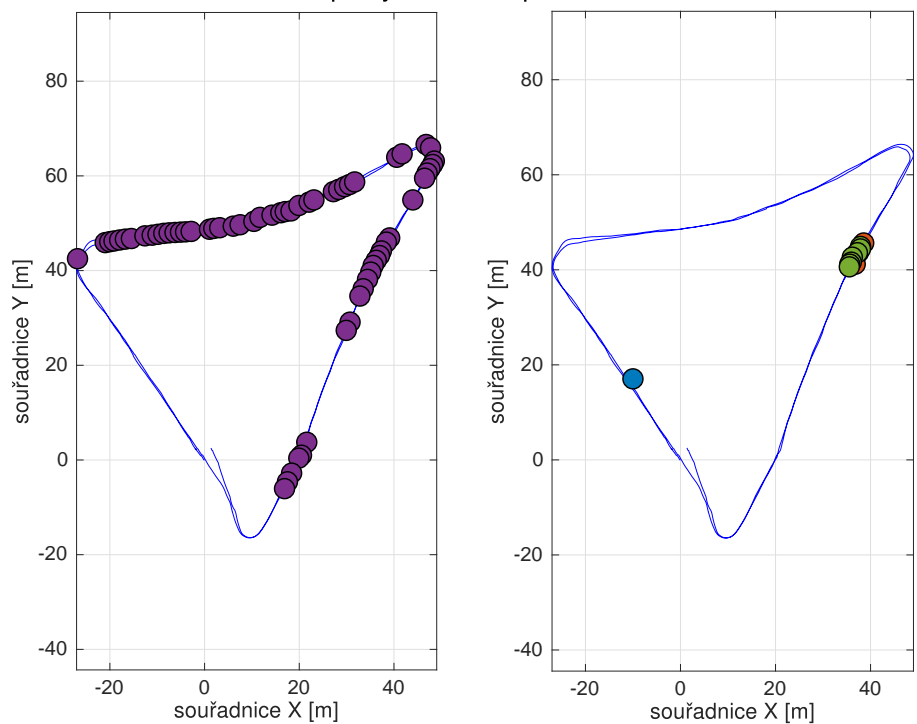
Obrázek 7.19: Lokalizace na mapě vytvořené 21. 2. 2019.

Lokalizace v mapě vytvořené dopoledne dne 7. 3. 2019



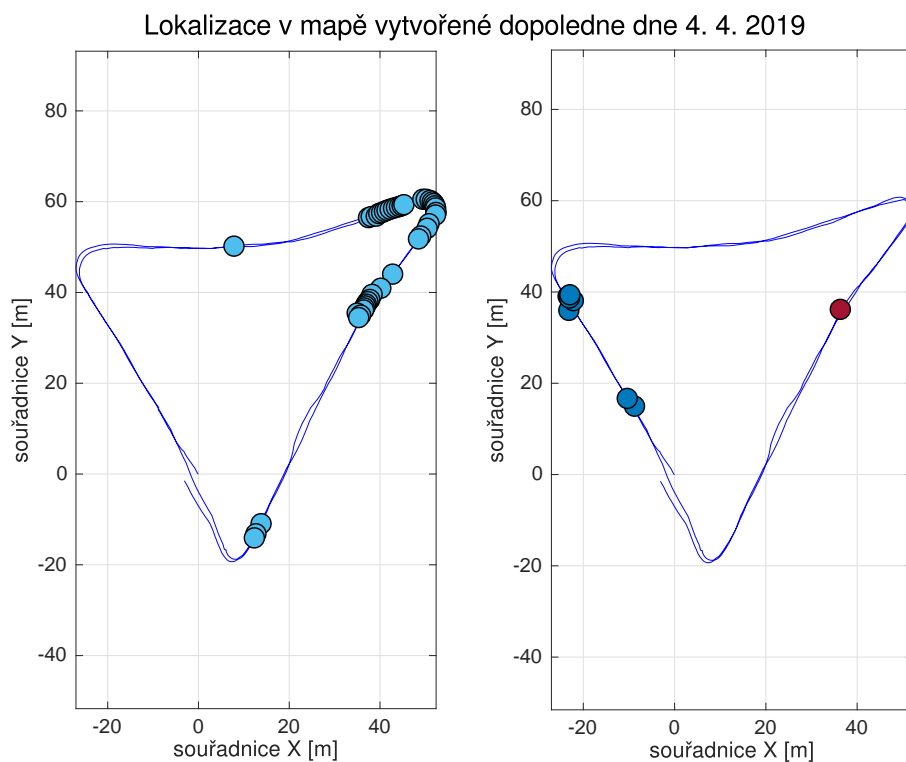
Obrázek 7.20: Lokalizace na mapě vytvořené 7. 3. 2019.

Lokalizace v mapě vytvořené dopoledne dne 21. 3. 2019

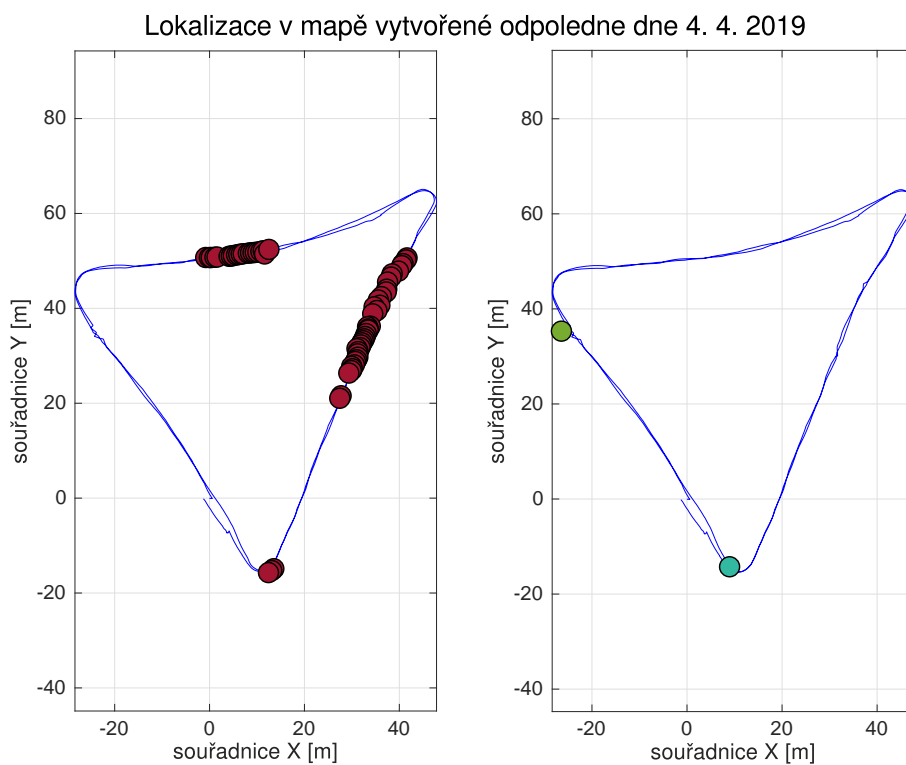


Obrázek 7.21: Lokalizace na mapě vytvořené 21. 3. 2019.

7.3. LOKALIZACE

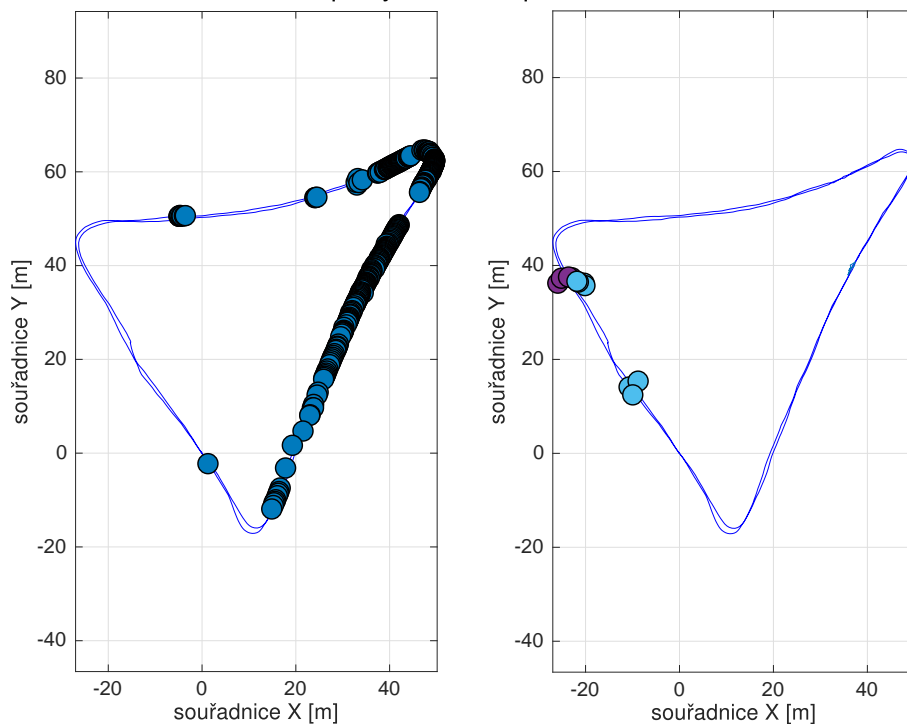


Obrázek 7.22: Lokalizace na mapě vytvořené dopoledne 4. 4. 2019.



Obrázek 7.23: Lokalizace na mapě vytvořené odpoledne 4. 4. 2019.

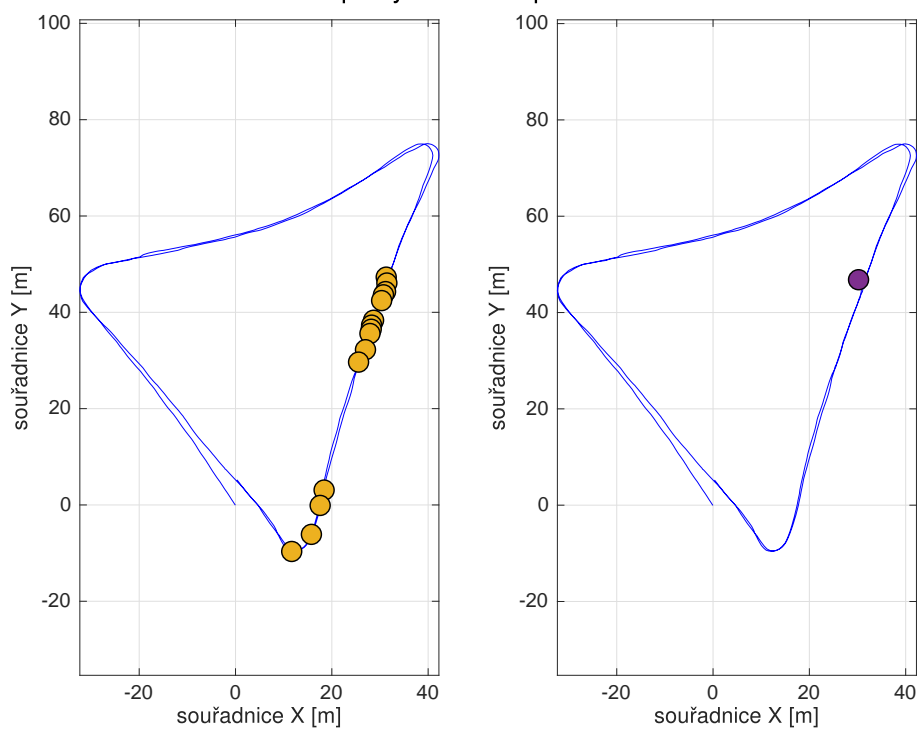
Lokalizace v mapě vytvořené dopoledne dne 11. 4. 2019



Obrázek 7.24: Lokalizace na mapě vytvořené dopoledne 11. 4. 2019.

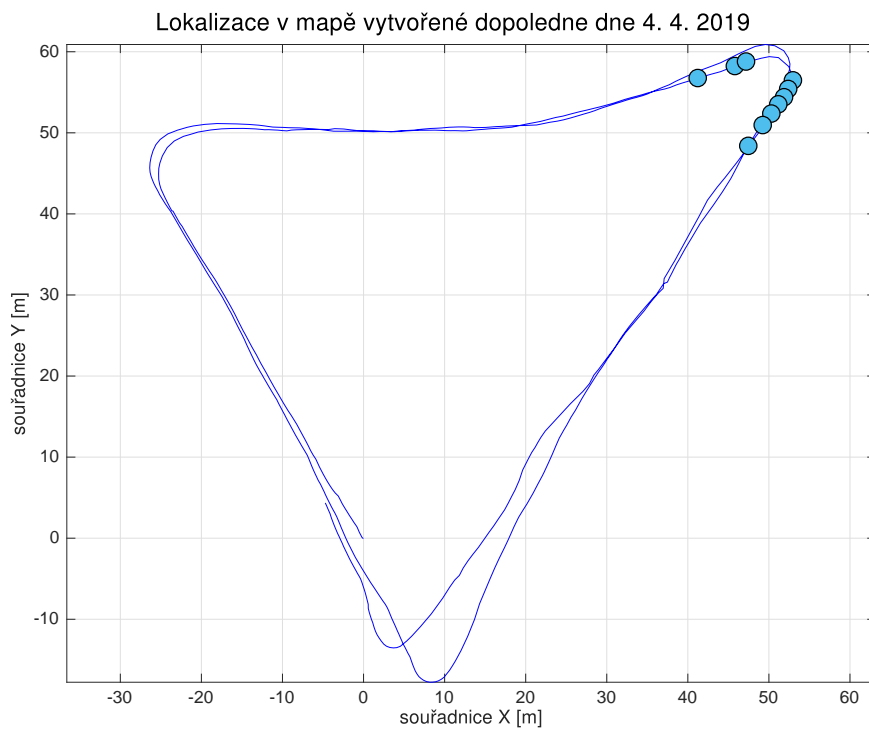
7.3.2. Lokalizace za použití metody SIFT

Lokalizace v mapě vytvořené dopoledne dne 7. 3. 2019

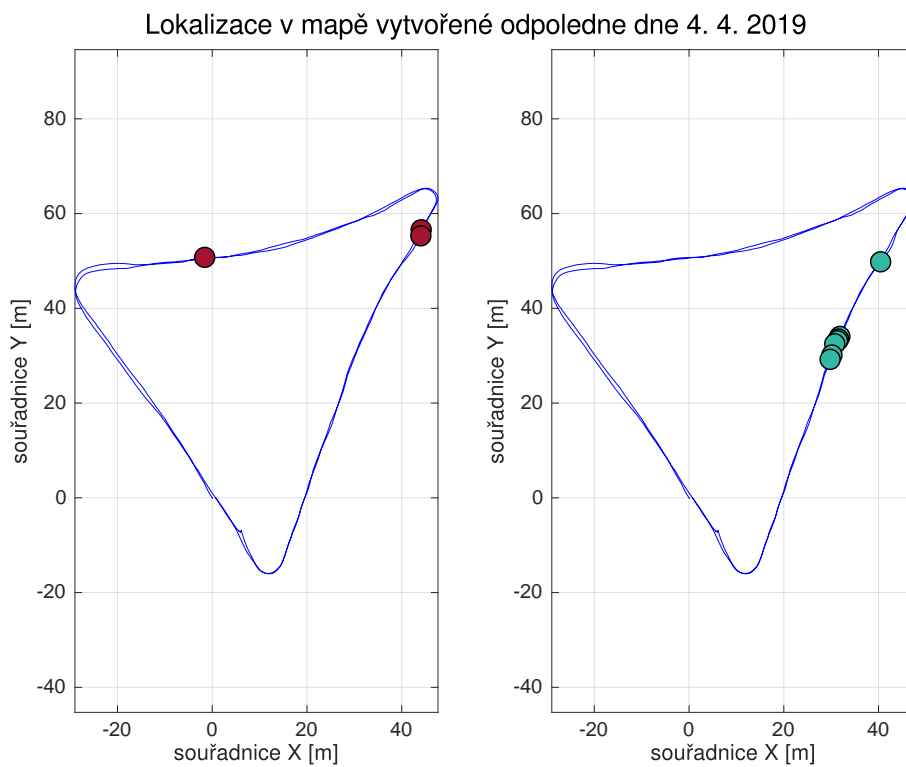


Obrázek 7.25: Lokalizace na mapě vytvořené 7. 3. 2019.

7.3. LOKALIZACE



Obrázek 7.26: Lokalizace na mapě vytvořené dopoledne 4. 4. 2019.



Obrázek 7.27: Lokalizace na mapě vytvořené odpoledne 4. 4. 2019.

7.3.3. Srovnání lokalizace

Testování lokalizace se odvíjelo i samotné použití dvou detektorů a deskriptorů SIFT a BRISK. Z obou provedených analýz vyplývá, že lokalizace ve venkovním prostředí funguje nespolehlivě. U metody BRISK je počet bodů lokalizace vyšší než u metody SIFT. V případě jakékoliv změny prostředí počet lokalizací klesá.

Detektor a deskriptor kontrastních objektů SIFT funguje ve venkovním prostředí méně než rohový detektor BRISK a je citlivější i na menší změny prostředí. To vyplývá i z výsledků, kde je patrné, že i z měření, která byla pořízena chvíli po mapování, je počet lokalizací daleko nižší oproti BRISK anebo ve většině případů nedošlo k žádné lokalizaci. Počet nalezených lokalizací je uveden v následujících tabulkách 7.1 a 7.2.

mapa \ data z měření	22. 11. 2018	21. 2. 2019	7. 3. 2019	21. 3. 2019 (dopoledne)	21. 3. 2019 (odpoledne)	4. 4. 2019 (dopoledne)	4. 4. 2019 (odpoledne)	11. 4. 2019 (dopoledne)	11. 4. 2019 (odpoledne)
22. 11. 2018	76	0	0	0	0	0	0	0	0
21. 2. 2019	0	77	8	4	4	0	0	0	3
7. 3. 2019	4	5	127	5	1	0	10	11	18
21. 3. 2019	0	5	0	63	7	0	0	1	0
4. 4. 2019 (dopoledne)	0	0	0	1	0	42	1	6	0
4. 4. 2019 (odpoledne)	0	0	0	0	1	0	64	0	1
11. 4. 2019 (dopoledne)	0	0	11	4	0	7	0	258	0

Tabulka 7.1: Tabulka s počtem správně lokalizovaných bodů na vytvořených mapách za použití detektoru a deskriptoru BRISK

mapa \ data z měření	22. 11. 2018	21. 2. 2019	7. 3. 2019	21. 3. 2019 (dopoledne)	21. 3. 2019 (odpoledne)	4. 4. 2019 (dopoledne)	4. 4. 2019 (odpoledne)	11. 4. 2019 (dopoledne)	11. 4. 2019 (odpoledne)
22. 11. 2018	-	-	-	-	-	-	-	-	-
21. 2. 2019	-	-	-	-	-	-	-	-	-
7. 3. 2019	0	0	15	1	0	0	0	0	0
21. 3. 2019	0	5	0	0	0	0	0	0	0
4. 4. 2019 (dopoledne)	0	0	0	0	0	0	10	0	0
4. 4. 2019 (odpoledne)	0	0	0	0	0	0	3	0	7
11. 4. 2019 (dopoledne)	0	0	0	0	0	7	0	0	0

Tabulka 7.2: Tabulka s počtem správně lokalizovaných bodů na vytvořených mapách za použití detektoru a deskriptoru SIFT

8. Závěr

V diplomové práci byla testována stereokamera ZED ve venkovním prostředí. Jejím cílem bylo analyzovat její použití v mobilní robotice. Testování probíhalo v první řadě ve výběru nastavení optimálních parametrů nebo metod. Ty jsou probrány v předchozích kapitolách. V dalším kroku bylo zapotřebí analyzovat mapování a lokalizaci za různých podmínek (osvětlení, rychlost pohybu kamery, statické a pohyblivé překážky). Data byla sesbírána offline za různých podmínek prostředí (náhlá změna osvětlení, statické a dynamické překážky, popřípadě vibrace způsobené nerovnostmi povrchu). Trajektorie byla měřena kolovou odometrií a pomocí GPS dat. V mapování se lokální kartézské souřadnicové systémy převedly do GPS souřadnic. Veškerá data byla porovnána vůči mapě, kde dané měření probíhalo. Poté se v případě SLAM testovala samotná lokalizace pro dva typy detektorů a deskriptorů pro detekci kontrastních objektů SIFT a pro detekci rohů BRISK. Počty lokalizovaných bodů při použití metody BRISK a SIFT byly v rámci analýzy lokalizace porovnány.

Ze samostatných analýz mapování vyplývá, že vizuální odometrie ZEDfu funguje spolehlivěji než měření GPS souřadnic nebo kolová odometrie. Ještě přesnější je použití SLAM mapování, kam vstupují data ze ZEDfu odometrie a dochází ke korekci trajektorie pomocí detekce uzavřených smyček. Samotná přesnost také závisí na použití detektorů a deskriptorů. U deskriptoru a detektoru kontrastních objektů SIFT docházelo k horší korekci mapy než u rohového detektoru a deskriptoru BRISK. Při lepších podmínkách osvětlení vycházely výsledky takřka totožné jak u SLAM, tak i u samotné ZEDfu vizuální odometrie.

V případě lokalizace, která se testovala pouze u SLAM, se výsledky zásadně lišily. Zatímco detektor BRISK nebyl náchylný na menší změny prostředí a také se vždy dokázal lokalizovat ihned po mapování prostředí, u SIFT i při menších změnách prostředí fungovala lokalizace velmi nespolehlivě. Kromě toho u tohoto detektoru a deskriptoru byla výpočtová náročnost vyšší a tím i pak pomalejší.

Z analýz vyplývá, že ZEDfu vizuální odometrie se SLAM mapování patří k přesným metodám. V případě lokalizace v SLAM se spolehlivost odvíjí od použitých detektorů a deskriptorů. Metoda BRISK vykazovala lepší výsledky jak u mapování, tak i lokalizace než za použití metody SIFT. Analýzy dle zadání byly úspěšně provedeny a zadané cíle v této diplomové práci byly tak splněny.

Seznam obrázků

2.1	Projekční matice kamery [8]	7
2.2	Znázornění perspektivní projekce [9]	8
2.3	Nákres epipolární geometrie [8]	9
2.4	Esenciální a fundametrální matice [11]	10
2.5	Perspektivní model kamery [12]	10
2.6	(a) Znázornění parametru f [15], (b) Znázornění parametru c [15], (c) Znázornění vlivu parametru s [15]	12
3.1	Blokový diagram vizuální odometrie [1]	14
3.2	Epipolární geometrie pro metodu z 2D do 2D [1]	15
4.1	Rozdělení do oktáv [21]	21
4.2	Znázornění deskriptorů pro metodu SIFT [22]	22
4.3	Ukázka detekce rohů na snímku [24].	23
4.4	Znázornění měřítkového prostoru metody BRISK [25]	23
4.5	Vzorkování okolo keypointu v metodě BRISK [25]	24
6.1	Blokový diagram RTAB-Map v systému ROS [26]	27
6.2	Blokový diagram SLAM v RTAB-Map při použití stereokamery [26]	28
6.3	Blokový diagram SLAM v RTAB-Map při vytváření lokální mapy [26]	29
6.4	Blokový diagram SLAM v RTAB-Map při vytváření globální mapy [26]	31
7.1	Rozměry stereokamery ZED [40]	32
7.2	Mapování ze dne 22. 11. 2018. GPS data (světle modrá), kolová odometrie (světle fialová).	35
7.3	Mapování ze dne 22. 11. 2018. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).	35
7.4	Mapování ze dne 21. 2. 2019. GPS data (světle modrá), kolová odometrie (světle fialová).	36
7.5	Mapování ze dne 21. 2. 2019. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).	36
7.6	Mapování ze dne 7. 3. 2019. GPS data (světle modrá), kolová odometrie (světle fialová).	37
7.7	Mapování ze dne 7. 3. 2019. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).	37
7.8	Mapování ze dne 21. 3. 2019. GPS data (světle modrá), kolová odometrie (světle fialová).	38
7.9	Mapování ze dne 21. 3. 2019. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).	38
7.10	Mapování ze dne 4. 4. 2019 v dopoledních hodinách. GPS data (světle modrá), kolová odometrie (světle fialová).	39
7.11	Mapování ze dne 4. 4. 2019 v dopoledních hodinách. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).	39
7.12	Mapování ze dne 4. 4. 2019 v odpoledních hodinách. GPS data (světle modrá), kolová odometrie (světle fialová).	40
7.13	Mapování ze dne 4. 4. 2019 v odpoledních hodinách. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).	40

7.14	Mapování ze dne 11. 4. 2019 v dopoledních hodinách. GPS data (světle modrá), kolová odometrie (světle fialová).	41
7.15	Mapování ze dne 11. 4. 2019 v dopoledních hodinách. ZEDfu (modrá), SLAM při použití metody BRISK (červená), SLAM při použití metody SIFT (žlutá).	41
7.16	Legenda pro grafy lokalizace.	43
7.17	Ukázky z daných měření	43
7.18	Lokalizace na mapě vytvořené 22. 11. 2018.	44
7.19	Lokalizace na mapě vytvořené 21. 2. 2019.	44
7.20	Lokalizace na mapě vytvořené 7. 3. 2019.	45
7.21	Lokalizace na mapě vytvořené 21. 3. 2019.	45
7.22	Lokalizace na mapě vytvořené dopoledne 4. 4. 2019.	46
7.23	Lokalizace na mapě vytvořené odpoledne 4. 4. 2019.	46
7.24	Lokalizace na mapě vytvořené dopoledne 11. 4. 2019.	47
7.25	Lokalizace na mapě vytvořené 7. 3. 2019.	47
7.26	Lokalizace na mapě vytvořené dopoledne 4. 4. 2019.	48
7.27	Lokalizace na mapě vytvořené odpoledne 4. 4. 2019.	48

Seznam tabulek

7.1	Tabulka s počtem správně lokalizovaných bodů na vytvořených mapách za použití detektoru a deskriptoru BRISK	50
7.2	Tabulka s počtem správně lokalizovaných bodů na vytvořených mapách za použití detektoru a deskriptoru SIFT	50

Literatura

1. SCARAMUZZA, D.; FRAUNDORFER, F. Visual Odometry [Tutorial]. *IEEE Robotics Automation Magazine*. 2011, roč. 18, č. 4, s. 80–92. ISSN 1070-9932. Dostupné z DOI: 10.1109/MRA.2011.943233.
2. MORAVEC, Hans Peter. *Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover*. Stanford, CA, USA: Stanford University, 1980. Disertační práce. AAI8024717.
3. MATTHIES, L.; SHAFER, S. Error modeling in stereo navigation. *IEEE Journal on Robotics and Automation*. 1987, roč. 3, č. 3, s. 239–248. ISSN 0882-4967. Dostupné z DOI: 10.1109/JRA.1987.1087097.
4. MATTHIES, Larry Henry. *Dynamic Stereo Vision*. Pittsburgh, PA, USA: Carnegie Mellon University, 1989. Disertační práce. AAI9023429.
5. LACROIX, S.; MALLET, A.; CHATILA, R. Rover self localization in planetary-like environments. In: *5th International Symposium on Artificial Intelligence, Robotics and Automation in Space, Noordwijk (The Netherlands)*. 1999, s. 433–440.
6. OLSON, C. F.; MATTHIES, L. H.; SCHOPPERS, H.; MAIMONE, M. W. Robust stereo ego-motion for long distance navigation. In: *Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No.PR00662)*. 2000, sv. 2, 453–458 vol.2. ISSN 1063-6919. Dostupné z DOI: 10.1109/CVPR.2000.854879.
7. NISTER, D.; NARODITSKY, O.; BERGEN, J. Visual odometry. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004*. 2004, sv. 1, s. I–I. ISSN 1063-6919. Dostupné z DOI: 10.1109/CVPR.2004.1315094.
8. VORÁČOVÁ, Šárka. Aplikace epipolární geometrie. In: *Sborník 24. konference Geometrie a počítačové grafiky*. Barborka, 2004, s. 236–242.
9. KITANI, Kris. 8. 2. Camera Matrix. *Carnegie Mellon University, The Robotics Institute* [online]. 2019-3-8 [cit. 2019-02-04]. Dostupné z: http://www.cs.cmu.edu/~16385/s17/Slides/11.1_Camera_matrix.pdf.
10. LONGUET-HIGGINS, H. C. Readings in Computer Vision: Issues, Problems, Principles, and Paradigms. In: FISCHLER, Martin A.; FIRSCHEIN, Oscar (ed.). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1987, kap. A Computer Algorithm for Reconstructing a Scene from Two Projections, s. 61–62. ISBN 0-934613-33-8. Dostupné také z: <http://dl.acm.org/citation.cfm?id=33517.33523>.
11. OPSAHL, Thomas. Lecture 7.2 Triangulation. *UiO Universitetet i Oslo* [online]. 2016-3-7 [cit. 2019-02-04]. Dostupné z: https://www.uio.no/studier/emner/matnat/its/UNIK4690/v16/forelesninger/lecture_7_2-triangulation.pdf.
12. *Camera Calibration and 3D Reconstruction* [online] [cit. 2019-02-04]. Dostupné z: https://docs.opencv.org/2.4/_images/pinhole_camera_model.png.
13. HARTLEY, Richard; ZISSERMAN, Andrew. *Multiple View Geometry in Computer Vision*. 2. vyd. New York, NY, USA: Cambridge University Press, 2003. ISBN 0521540518.

LITERATURA

14. PIVOŇKA, Tomáš. *Vizuální odometrie pro dynamickou rekonstrukci obrazu*. Praha, 2018. Diplomová práce. České vysoké učení technické v Praze. Výpočetní a informační centrum. Vedoucí práce Přeucil Libor.
15. HAAVARDSHOLM, Trym. Lecture 1.4. The perspective camera model. *UiO Universitetet i Oslo* [online]. 2019-1-22 [cit. 2019-02-04]. Dostupné z: https://www.uio.no/studier/emner/matnat/its/UNIK4690/v16/forelesninger/lecture_1_4-the-perspective-camera-model.pdf.
16. NISTÉR, David. An Efficient Solution to the Five-Point Relative Pose Problem. *IEEE Trans. Pattern Anal. Mach. Intell.* 2004, roč. 26, č. 6, s. 756–777. ISSN 0162-8828. Dostupné z DOI: 10.1109/TPAMI.2004.17.
17. LONGUET-HIGGINS, H. C. Readings in Computer Vision: Issues, Problems, Principles, and Paradigms. In: FISCHLER, Martin A.; FIRSCHEIN, Oscar (ed.). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 1987, kap. A Computer Algorithm for Reconstructing a Scene from Two Projections, s. 61–62. ISBN 0-934613-33-8. Dostupné také z: <http://dl.acm.org/citation.cfm?id=33517.33523>.
18. ARUN, K. S.; HUANG, T. S.; BLOSTEIN, S. D. Least-Squares Fitting of Two 3-D Point Sets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1987, roč. PAMI-9, č. 5, s. 698–700. ISSN 0162-8828. Dostupné z DOI: 10.1109/TPAMI.1987.4767965.
19. MAIMONE, Mark W.; CHENG, Yang; MATTHIES, Larry H. Two years of Visual Odometry on the Mars Exploration Rovers. *J. Field Robotics*. 2007, roč. 24, s. 169–186.
20. TAREEN, S. A. K.; SALEEM, Z. A comparative analysis of SIFT, SURF, KAZE, AKAZE, ORB, and BRISK. In: *2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)*. 2018, s. 1–10. Dostupné z DOI: 10.1109/ICOMET.2018.8346440.
21. LOWE, David G. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*. 2004, roč. 60, č. 2, s. 91–110.
22. *A Short introduction to descriptors* [online] [cit. 2019-04-23]. Dostupné z: <https://gilscvblog.com/2013/08/18/a-short-introduction-to-descriptors/>.
23. *Feature Detection and Description* [online] [cit. 2019-04-17]. Dostupné z: https://docs.opencv.org/3.0-beta/doc/py_tutorials/py_feature2d/py_table_of_contents_feature2d/py_table_of_contents_feature2d.html.
24. ROSTEN, E.; PORTER, R.; DRUMMOND, T. Faster and Better: A Machine Learning Approach to Corner Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2010, roč. 32, č. 1, s. 105–119. ISSN 0162-8828. Dostupné z DOI: 10.1109/TPAMI.2008.275.
25. LEUTENEGGER, Stefan; CHLI, Margarita; SIEGWART, Roland Y. BRISK: Binary Robust Invariant Scalable Keypoints. In: *Proceedings of the 2011 International Conference on Computer Vision*. Washington, DC, USA: IEEE Computer Society, 2011, s. 2548–2555. ICCV '11. ISBN 978-1-4577-1101-5. Dostupné z DOI: 10.1109/ICCV.2011.6126542.

26. LABBÉ, Mathieu; MICHAUD, François. RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *Journal of Field Robotics*. 2019, roč. 36, č. 2, s. 416–446. Dostupné z DOI: 10.1002/rob.21831.
27. LABBÉ, Mathieu; MICHAUD, François. RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *Journal of Field Robotics*. 2019, roč. 36, č. 2, s. 416–446. Dostupné z DOI: 10.1002/rob.21831.
28. LABBÉ, Mathieu; MICHAUD, François. Long-term online multi-session graph-based SPLAM with memory management. *Autonomous Robots*. 2018, roč. 42, č. 6, s. 1133–1150. ISSN 1573-7527. Dostupné z DOI: 10.1007/s10514-017-9682-5.
29. LABBÉ, M.; MICHAUD, F. Online global loop closure detection for large-scale multi-session graph-based SLAM. In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2014, s. 2661–2666. ISSN 2153-0858. Dostupné z DOI: 10.1109/IR0S.2014.6942926.
30. LABBÉ, M.; MICHAUD, F. Appearance-Based Loop Closure Detection for On-line Large-Scale and Long-Term Operation. *IEEE Transactions on Robotics*. 2013, roč. 29, č. 3, s. 734–745. ISSN 1552-3098. Dostupné z DOI: 10.1109/TR0.2013.2242375.
31. LABBÉ, M.; MICHAUD, F. Memory management for real-time appearance-based loop closure detection. In: *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2011, s. 1271–1276. ISSN 2153-0866. Dostupné z DOI: 10.1109/IR0S.2011.6094602.
32. QUIGLEY, Morgan; CONLEY, Ken; GERKEY, Brian P.; FAUST, Josh; FOOTE, Tully; LEIBS, Jeremy; WHEELER, Rob; NG, Andrew Y. ROS: an open-source Robot Operating System. In: *ICRA Workshop on Open Source Software*. 2009.
33. KONOLIGE, Kurt. Small Vision Systems: Hardware and Implementation. In: SHIRAI, Yoshiaki; HIROSE, Shigeo (ed.). *Robotics Research*. London: Springer London, 1998, s. 203–212. ISBN 978-1-4471-1580-9.
34. RUSU, R. B.; COUSINS, S. 3D is here: Point Cloud Library (PCL). In: *2011 IEEE International Conference on Robotics and Automation*. 2011, s. 1–4. ISSN 1050-4729. Dostupné z DOI: 10.1109/ICRA.2011.5980567.
35. LABBÉ, M.; MICHAUD, F. Online global loop closure detection for large-scale multi-session graph-based SLAM. In: *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2014, s. 2661–2666. ISSN 2153-0858. Dostupné z DOI: 10.1109/IR0S.2014.6942926.
36. LABBÉ, Mathieu; MICHAUD, François. Long-term online multi-session graph-based SPLAM with memory management. *Autonomous Robots*. 2018, roč. 42, č. 6, s. 1133–1150. ISSN 1573-7527. Dostupné z DOI: 10.1007/s10514-017-9682-5.
37. GRISETTI, G.; KUMMERLE, R.; STACHNISS, C.; BURGARD, W. A Tutorial on Graph-Based SLAM. *IEEE Intelligent Transportation Systems Magazine*. 2010, roč. 2, č. 4, s. 31–43. ISSN 1939-1390. Dostupné z DOI: 10.1109/MITS.2010.939925.

LITERATURA

38. KÜMMERLE, R.; GRISETTI, G.; STRASDAT, H.; KONOLIGE, K.; BURGARD, W. G2o: A general framework for graph optimization. In: *2011 IEEE International Conference on Robotics and Automation*. 2011, s. 3607–3613. ISSN 1050-4729. Dostupné z DOI: 10.1109/ICRA.2011.5979949.
39. DELLAERT, Frank. *Factor graphs and GTSAM: A hands-on introduction*. 2012. Technická zpráva. Georgia Institute of Technology.
40. *ZED Stereo Camera - Stereolabs* [online] [cit. 2019-04-29]. Dostupné z: <https://www.stereolabs.com/zed/>.
41. BRADSKI, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*. 2000.
42. ORTIZ, Luis; CABRERA, Elizabeth; GONÇALVES, Luiz. Depth Data Error Modeling of the ZED 3D Vision Sensor from Stereolabs. *Electronic Letters on Computer Vision and Image Analysis*. 2018, roč. 17. Dostupné z DOI: 10.5565/rev/elcvia.1084.